

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



Seguimento Multissensorial de Objetos Dinâmicos para Aplicações Robóticas

Miguel Alexandre Vilela Bertão

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Orientador: Doutor Andry Maykol Gomes Pinto

Co-orientador: Doutor Aníbal Castilho Coimbra de Matos

25 de Junho de 2018

Resumo

A utilização de veículos autónomos de superfície tem sofrido uma crescente procura em diversas áreas, nomeadamente na realização de tarefas potencialmente perigosas para o Ser Humano, tendo cada vez mais influência na monitorização do ambiente circundante. Para garantir a evolução destes veículos, é necessário dotá-los da capacidade de manterem uma perceção constante de tudo o que passa à sua volta, em específico, dos movimentos dos objetos circundantes. A par desta procura, verificam-se também desenvolvimentos nas tecnologias e métodos de aquisição sensorial que, para além de exibirem preços apelativos, apresentam resultados positivos nas operações a curta e longa distância. Esta melhoria não implica, todavia, uma estagnação no desenvolvimento, sendo necessário continuar a estudar e a aperfeiçoar métodos e sistemas capazes de utilizar as mais variadas tecnologias disponíveis para realizar estas tarefas, permitindo o alargamento da gama de ambientes e condições a que estes sistemas autónomos podem operar. Através da aplicação de métodos de seguimento, é possível a perceção dos movimentos dos objetos circundantes. Deste modo, este tipo de sistemas consegue manter, constantemente, a perceção das posições desses objetos e estimar os seus movimentos futuros à semelhança do Homem.

Assim, a presente dissertação passou inicialmente pela escolha de um método de deteção de visão monocular para a identificação dos objetos no mundo. Posteriormente, é explicitado o desenvolvimento de uma abordagem de seguimento baseado na fusão sensorial de tecnologias distintas e na aplicação de filtros de Kalman para estimação de movimentos. Tudo isto, para que seja possível, futuramente, a sua aplicação em contextos reais de operação. Por fim, são descritos os testes realizados em cenários exteriores e interiores, onde foi avaliado o desempenho e fatores condicionantes deste sistema.

Das experiências realizadas, verificou-se que o seguimento resultante de fusão sensorial apresenta resultados positivos e constitui uma melhoria passível de ser implementada em veículos autónomos.

Abstract

The use of autonomous surface vehicles has been going through an increase of demand in various areas, namely in the execution of potentially dangerous tasks to the human being, having a great influence in the monitoring of the surrounding environment. To guarantee the evolution of these vehicles, is necessary to endow them with the ability to maintain a constant perception of everything happening around them, specifically, the movements of surrounding objects. Along with this demand, developments have been occurring in the technologies and methods of sensory acquisition that, besides exhibiting appealing prices, also present positive results in operations of short and long range. However, these improvements do not imply that development can become stagnant; it is necessary to continue the study and development of methods and systems capable of using the available technologies to perform these tasks, allowing the increase of environments and conditions on which these autonomous systems can function. The perception of movement of the surrounding objects is possible through the application of tracking methods. Thus, this type of systems can maintain, constantly, the perception of the positions of these objects and estimate their future movements, in the likeness of man.

Firstly, in this dissertation, is chosen a monocular vision method of detection to identify objects in the world. Afterwards, is explicated the development of a tracking approach based on the sensory fusion of distinct technologies and on the application of Kalman filters to estimate movements. The goal is that, in the future, it will be possible to apply it in a real operation context. Lastly, the tests performed in outdoors and indoors scenarios are described, where the performance and conditioning factors were evaluated.

From the accomplished experiments, it was verified that the tracking resulting from sensory fusion presents positive results and constitutes an improvement that can be applied to autonomous surface vehicles.

Agradecimentos

Esta dissertação representa o culminar de um ciclo muito importante na minha vida que, embora tenha sido uma tarefa exigente e árdua, me proporcionou uma enorme satisfação em realizá-la. Contudo, esta teria sido impossível de alcançar sozinho.

Agradeço, desde já, ao meu orientador, Professor Andry Pinto, por todo o apoio, exigência e paciência que teve comigo e, acima de tudo, pelos ensinamentos, palavras de incentivo e oportunidades que me propiciou ao longo desta etapa, tornando possível atingir este resultado. Agradeço, ainda, à equipa do CRAS por todo o apoio, em especial, à Rita e à Alexandra, pela disponibilidade, prontidão e ajuda dadas, independentemente da dimensão do problema, e, acima de tudo, pela boa disposição demonstrada, que ajudou a reduzir o *stress* que esta fase acarreta. Por isso, muito obrigado a todos!

Quero agradecer aos meus pais, pelo apoio incondicional que me deram, não só na realização da minha vida académica, porque sempre acreditaram em mim, mas também nestes 23 anos de vida, onde estiveram presentes e que fizeram de tudo para que nada faltasse. A vós, o meu muito obrigado. À minha irmã que seguiu este trajeto atentamente e sempre acreditou em mim. Aos meus avós, pelos mimos e confiança que sempre depositaram em mim.

Ao João, ao Rui e ao Tiago, pelo companheirismo nesta caminhada de 5 anos, em que juntos lutámos pela concretização deste objetivo. É nos momentos de aperto em que vemos quem são as pessoas com quem podemos contar e a quem podemos chamar de amigos. Acima de tudo, agradeço-vos por serem vocês mesmos, porque a realidade é que, se estes últimos anos fossem passados apenas com pessoas normais, teriam sido um aborrecimento. Obrigado pela amizade!

E, como não podia deixar de ser, quero agradecer, especialmente, à Beatriz, a quem a concretização deste objetivo mais se deve. Obrigado por estares presente e seres o meu porto de abrigo. Por sempre acreditares em mim e me apoiares incondicionalmente, por toda a paciência que tiveste comigo (e eu sei que é preciso muita) e por toda a preocupação e carinho que sempre demonstraste, obrigado.

A todos, obrigado por tudo!

Miguel Bertão

“The only man who never makes mistakes is the man who never does anything.”

Theodore Roosevelt

Conteúdo

Agradecimentos	v
1 Introdução	1
1.1 Contexto	1
1.2 Motivação	2
1.3 Objetivos	3
1.4 Estrutura da Dissertação	3
2 Revisão Bibliográfica	5
2.1 Introdução	5
2.2 Seguimento de objetos com recurso a tecnologias passivas	6
2.3 Seguimento de objetos com recurso a tecnologias ativas	9
2.4 Seguimento de objetos com recurso a tecnologias passivas e ativas	11
2.5 Conclusão	12
3 Seguimento Multissensorial	13
3.1 Introdução	13
3.2 Arquitetura sensorial	14
3.2.1 Sensores do Sistema de Aquisição	16
3.3 Algoritmo de Seguimento	18
3.3.1 Introdução	18
3.3.2 Metodologia e Especificações Técnicas	20
3.4 Resultados	34
3.4.1 Caso de Estudo	34
3.4.2 Conclusão	43
4 Conclusão e Trabalho Futuro	45
4.1 Conclusão	45
4.2 Satisfação dos Objetivos e Trabalho Futuro	46
Referências	47

Lista de Figuras

2.1	Diagrama de fluxo representativo do sistema de detecção e seguimento baseado em visão [1].	7
2.2	Imagem de um caso de teste com boas condições de visibilidade. A seta vermelha indica a detecção e presença de uma aeronave intrusa, a uma distância de cerca de 1.3 km. [1]	8
2.3	Imagem de um caso de teste com piores condições atmosféricas e de visibilidade do que as representadas na figura anterior. A seta vermelha indica a detecção e presença de uma aeronave intrusa, a uma distância de cerca de 1.3 km [1].	8
2.4	Deteção por LIDAR e a correspondente vista na imagem. Em 1: Carro ocluso. Em 2: Diversas formas de carros [2].	10
2.5	Através do algoritmo de <i>clustering</i> , foram determinados três alvos distintos[3].	11
2.6	Deteção e seguimento de múltiplos obstáculos. A verde estão representadas potenciais ameaças, enquanto que a vermelho é expressa a ameaça de colisão mais iminente [4].	12
3.1	Vista geral do sistema e das ferramentas utilizadas	14
3.2	Exemplo ilustrativo de uma interação entre 3 nós, no ROS	15
3.3	Câmara - Mako G-125 [5]	16
3.4	Sensor LIDAR - VLP-16 [6]	17
3.5	Enquadramento da dissertação	18
3.6	Fluxograma das etapas do algoritmo desenvolvido	19
3.7	Exemplos de marcadores [7]	20
3.8	Marcador ArUco	21
3.9	Exemplo de interação do Aruco com dois marcadores	22
3.10	Nuvem de Pontos do ambiente	23
3.11	Transformação e fusão dos dados sensoriais para um referencial comum	23
3.12	Discrepância Sensorial	24
3.13	Nuvem de pontos antes da segmentação	25
3.14	Região de interesse obtida após filtragem	26
3.15	Algoritmo de fusão sensorial	28
3.16	Output da fusão sensorial	29
3.17	Diagrama lógico do algoritmo de seguimento multi-objeto [8]	30
3.18	Caso de estudo em ambiente fechado 1	35
3.19	Caso de estudo em ambiente fechado 2	35
3.20	Caso de estudo em ambiente fechado 3	35
3.21	Caso de estudo em ambiente fechado 4	36
3.22	Caso de estudo em ambiente fechado 5	36
3.23	Caso de estudo em ambiente fechado 6	36

3.24	Caso de estudo em ambiente fechado 7	37
3.25	Caso de estudo em ambiente fechado 8	37
3.26	Percurso efetuado pelos objetos e o processo de estimação de cada filtro de Kalman em ambiente interior	38
3.27	Caso de estudo no exterior 1	39
3.28	Caso de estudo no exterior 2	40
3.29	Caso de estudo no exterior 3	40
3.30	Caso de estudo no exterior 4	40
3.31	Caso de estudo no exterior 5	41
3.32	Caso de estudo no exterior 6	41
3.33	Percurso efetuado pelos objetos e o processo de estimação de cada filtro de Kalman em ambiente exterior	42

Lista de Tabelas

3.1	Especificações do Mako G-125 [5]	16
3.2	Especificações do VLP-16 [6]	17

Abreviaturas e Símbolos

ASV	Autonomous Surface Vehicle
CRAS	Centro de Robótica e Sistemas Autónomos
FEUP	Faculdade de Engenharia da Universidade do Porto
LIDAR	LIght Detection And Ranging
NED	North East Down
RANSAC	RANdom SAmple Consensus
RADAR	RAdio Detection And Ranging
ROI	Region Of Interest
ROS	Robotic Operative System
SONAR	SOund Navigation And Ranging
TOF	Time of Flight

Capítulo 1

Introdução

1.1 Contexto

O uso dos mais variados tipos de veículos autónomos tem sofrido uma crescente procura, motivada pelas suas características e potencialidades, das quais o ser humano consegue tirar partido por executarem tarefas que podem constituir uma ameaça para o Homem.

O desenvolvimento de estruturas capazes de nos substituírem nas mais variadas tarefas foi, desde sempre, desejada. A evolução dos *Autonomous Surface Vehicles* (ASV), aliada ao desenvolvimento dos computadores, nomeadamente ao aumento da capacidade de processamento, memória e rapidez, permitiu a melhoria das aptidões operativas que estes veículos podem realizar. Atualmente, este tipo de veículos é utilizado na realização das mais variadas operações, como a vigilância de fronteiras e de território, por vezes hostil, monitorização do ambiente, fauna e flora, recolha de dados atmosféricos e missões de salvamento.

Para que seja possível a utilização dos ASV, é necessário que estes consigam navegar autonomamente. Para tal, é vital que os dados sensoriais sejam adquiridos com precisão, de forma a conseguir obter-se uma localização espacial fiável.

Durante o dia, a elevada reflexão da superfície do meio aquático marítimo, aliada à sua irregularidade (ondas) e à ausência de luz durante a noite, torna muito difícil a deteção de objetos através de câmaras. Por este motivo, são muitas vezes utilizadas técnicas de identificação visual com recurso a tecnologias infravermelhas como Light Detection And Ranging (LIDAR), onde o robô, recorrendo aos sensores nele acoplados, consegue criar um mapa do ambiente circundante, verificando ou refutando a existência de objetos ao seu redor.

Uma vez que o sistema que se pretende implementar tem como cenário de operação o ambiente marítimo, e sendo este um espaço aberto e à superfície, é possível utilizar um vasto leque de tecnologias. Neste meio, os sinais de rádio propagam-se até longas distâncias, tornando técnicas como o Radar e Sistema de Posicionamento Global (GPS) apelativas. Também é possível usar sensores baseados em acústica, mas dada a necessidade da existência de sistemas com elevada precisão e taxa de atualização, estes métodos não são normalmente utilizados, sendo preferível

o uso de tecnologias baseadas em LIDAR e visão para a detecção de objetos e ondas rádio para comunicações.

A tecnologia LIDAR tem sido amplamente usada como sistema de mapeamento ou como instrumento de detecção de objetos complementar, sendo empregue como uma ferramenta auxiliar, confirmando ou refutando possíveis casos positivos em sistemas de detecção, conferindo um maior grau de precisão aos mesmos. Contudo, devido ao elevado desenvolvimento que esta tem vindo a demonstrar, nomeadamente em sensores capazes de fazer mapeamento em 3D com milhões de pontos e com taxas de atualização elevadas, é possível criar sistemas autónomos capazes de utilizar LIDAR como método principal de detecção, localização e seguimento de objetos.

Neste sentido, esta dissertação visa estudar um sistema de seguimento (*tracking*) de pontos de interesse, com recurso à conjugação de múltiplos sensores, focando-se nas tecnologias de Lidar e imagem.

1.2 Motivação

O uso de sistemas robotizados tem vindo a ser procurado cada vez mais devido às suas características, que permitem, num universo de possibilidades cada vez maior, substituir o ser humano em tarefas que possam ser potencialmente perigosas à sua vida ou até mesmo porque o fazem rápida e incessantemente.

Uma aplicação do uso destes sistemas é o do patrulhamento do território em alto mar, onde o risco para a vida humana é grande, devido ao isolamento e às condições adversas que naquele local se podem encontrar. Para o robô conseguir realizar uma vigilância, este tem que possuir mecanismos que permitam detetar, identificar e seguir com bastante exatidão objetos que circulem à superfície da água, como é o caso de embarcações e naufragos. Uma forma de se conseguir obter esta capacidade, é através do uso de tecnologias baseadas em LIDAR, que utiliza a informação da distância do sensor a milhões de pontos no espaço para construir um mapa com elevada precisão do local onde se encontra e, assim, identificar possíveis objetos à superfície. Este método de detecção e seguimento apresenta uma maior exatidão e resultados superiores a métodos que utilizam imagem provenientes de câmaras para a mesma finalidade, já que estas possuem uma baixa resolução a médias distâncias, entre 30 a 80 metros, revelando, por vezes, dificuldades em detetar corpos.

Tendo em conta que cada técnica, *per si*, incorpora uma fragilidade, o uso associado de várias tecnologias pode anular ou inibir as eventuais vulnerabilidades de cada uma. Este princípio justifica o desenvolvimento de algoritmos que integram estas tecnologias, capazes de detetar, identificar e seguir objetos num ambiente próximo, juntamente com a sucessiva redução de custos que os componentes utilizados têm vindo a sofrer, torna-se possível a implementação destes módulos nos mais variados sistemas autónomos, tornando a exploração e desenvolvimento destes métodos relevante.

1.3 Objetivos

Esta dissertação pretende desenvolver um sistema de seguimento de pontos de interesse com recurso a múltiplos sensores, que possa ser aplicado a sistemas autónomos de superfície, em ambiente marinho.

Pressupõe a realização dos seguintes objetivos:

- Escolha do método e sistema de deteção de objetos;
- Estudo e desenvolvimento de um método de seguimento baseado na fusão de dados sensoriais provenientes de imagens e LIDAR, e em filtros de Kalman que permitam a estimação dos movimentos dos objetos;
- Realização de testes para validação de algoritmos e capacidade operativa do sistema; e
- Caracterização experimental das técnicas utilizadas para validação dos dados.

1.4 Estrutura da Dissertação

Para além da introdução, esta dissertação contém mais 2 capítulos. No capítulo 2, é descrito o estado da arte e são apresentados trabalhos relacionados. No capítulo 3, são evidenciadas e descritas as tecnologias e ferramentas utilizadas, e expostos os algoritmos desenvolvidos, assim como os seus resultados. No capítulo 4 é realizada uma apreciação do trabalho desenvolvido, e uma sugestão de trabalho futuro como processo de melhoria.

Capítulo 2

Revisão Bibliográfica

Serve o presente capítulo para expor o Estado da Arte referente a esta dissertação. Nele será realizada uma revisão bibliográfica, sendo descritas as tecnologias utilizadas no processo de seguimento, assim como os processos e sistemas que permitem o seu funcionamento.

2.1 Introdução

Qualquer que seja o sistema de seguimento existente, é-lhe inerente o recurso a um sensor capaz de detetar e fornecer a posição espacial de um ou mais objetos, permitindo a perceção do mundo circundante.

Existe um vasto leque de tecnologias que permitem obter informação sobre o meio em que se inserem. Os sensores, e as respetivas tecnologias utilizadas para realizar esta função, podem ser de dois tipos: ativos e passivos. Alguns dos sensores ativos, para a obtenção da posição relativa dos objetos, usam tecnologias do tipo *range finders*. Estas efetuam varrimentos do espaço circundante com lasers (LIDAR), som (SONAR), ou ondas rádio (RADAR), para adquirirem informações espaciais do meio, usando o princípio de *time-of-flight* (TOF). Ao contrário das tecnologias supra-mencionadas, que interagem ativamente com o meio para a recolha de informações, os sensores passivos reúnem estes dados através da recolha de informação que o meio exterior lhes fornece. Entre os exemplos mais conhecidos desta tecnologia, encontram-se os mais variados tipos de câmaras e os sensores infravermelhos.

Visto que existem inúmeras fontes de aquisição e deteção de objetos, a realização do seguimento dos mesmos pode ser realizada sobre uma, ou sobre a conjugação das mais variadas tecnologias anteriormente apresentadas. Assim sendo, este capítulo encontra-se dividido nas seguintes temáticas:

- Seguimento de objetos com recurso a tecnologias passivas;
- Seguimento de objetos com recurso a tecnologias ativas;
- Seguimento de objetos com recurso a tecnologias passivas e ativas.

2.2 Seguimento de objetos com recurso a tecnologias passivas

Atualmente, devido ao elevado desenvolvimento das tecnologias de visão, as câmaras de visão têm sido usadas como principal ferramenta, quando se fala do uso das tecnologias passivas na deteção de corpos. Graças ao reduzido custo destes sensores óticos, é possível a obtenção, com relativa facilidade, de câmaras com características e resoluções apelativas e de elevado desempenho. No entanto, a sua grande desvantagem reside no facto de estas serem facilmente afetadas pelas condições climáticas e de luminosidade. Em alto mar, por exemplo, chegam praticamente a perder quase totalmente a sua funcionalidade, perante condições de forte ou reduzida luminosidade.

Apesar das condições de operabilidade acima referidas, e graças à sua principal vantagem, foram desenvolvidos inúmeros sistemas de seguimento que assentam e dependem substancialmente neste tipo de sensores. Um desses tipos de sistemas desenvolvidos [1] consiste na deteção e seguimento de aeronaves de forma a prever e evitar colisões entre aeronaves em pleno voo. Em [1], os autores propõem uma combinação entre técnicas de processamento de imagem e algoritmos de estimativa de estado, como forma de detetar e estimar a posição de uma aeronave durante o voo. Após a confirmação de uma deteção, é gerado um *template* de obstáculo e o objeto é seguido nos *frames* subsequentes, através de algoritmos de correlação [9]. Assim que o seguimento visual é alcançado, e por sua vez a posição dos objetos é, então, utilizado um filtro de Kalman, com o intuito de obter estimativas de estado e velocidade. De seguida, é apresentado, na figura 2.1, o algoritmo descrito anteriormente, sob forma de diagrama de fluxo.

Com um algoritmo baseado em tecnologias passivas, nomeadamente a tecnologia de visão proveniente de câmaras, foi possível extrair informações de movimento do voo da avioneta, em particular, a posição e a elevação, em coordenadas NED e derivadas angulares. Apesar deste sistema apresentar pouca robustez em relação ao clima, que afeta invariavelmente as condições de luminosidade, foi capaz de detetar aeronaves a distâncias superiores a 1 quilómetro, distância à qual a aeronave se torna quase impercetível ao olho humano, como é possível averiguar nas figuras 2.2 e 2.3.

Em [10], é introduzido um sistema capaz de aprender, detetar e realizar o seguimento de objetos, sendo que, neste caso em particular, os objetos são jogadores de hóquei no gelo. Para isso, foram usados dois algoritmos, misturando o algoritmo em cascata AdaBoost, de Viola e Jones [11], com um filtro de partículas. Da união do modelo dinâmico previsto anteriormente, com as deteções de jogadores do *frame* anterior, proveniente do AdaBoost, é possível gerar uma distribuição capaz de detetar rapidamente jogadores que estão a entrar em cena, enquanto que a filtragem permite a manutenção do seguimento de jogadores isoladamente.

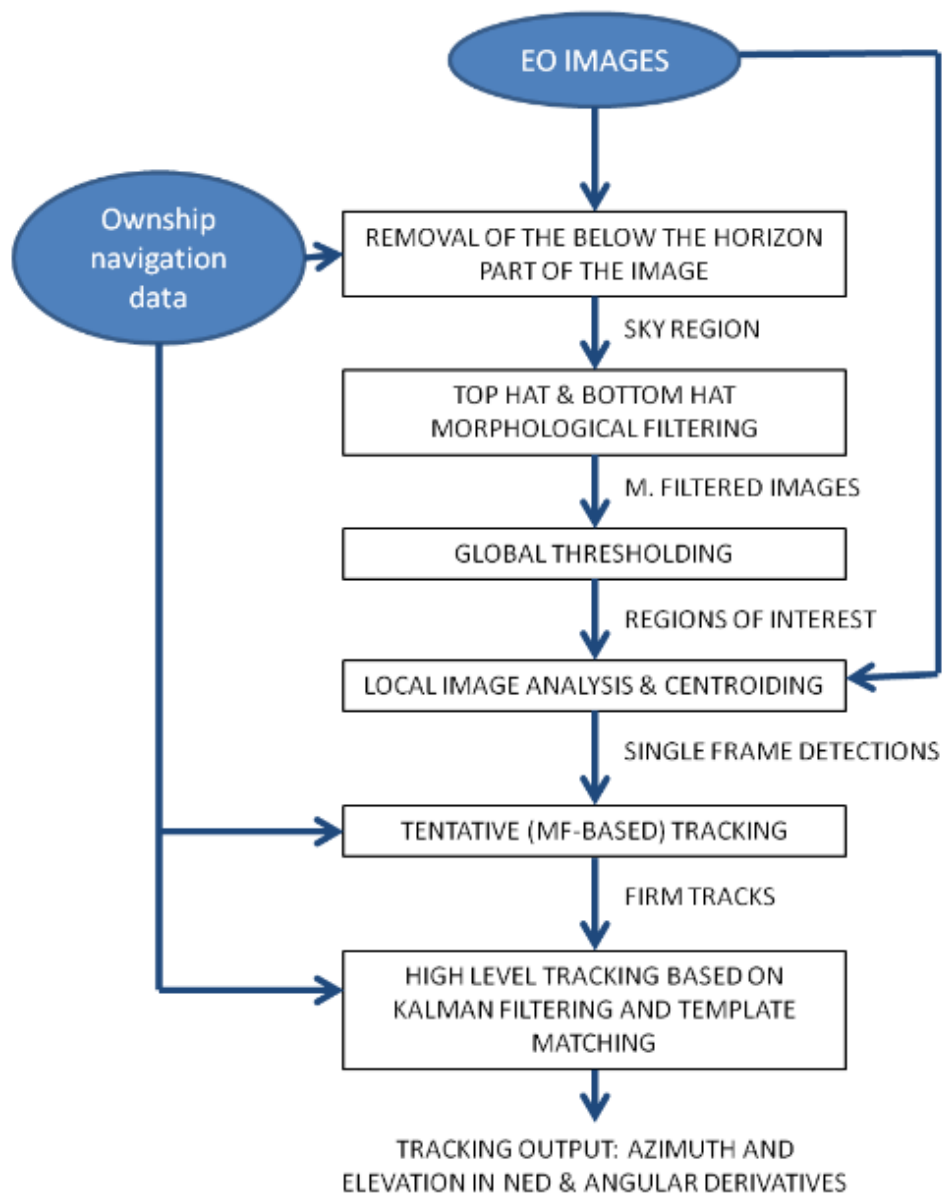


Figura 2.1: Diagrama de fluxo representativo do sistema de deteção e seguimento baseado em visão [1].



Figura 2.2: Imagem de um caso de teste com boas condições de visibilidade. A seta vermelha indica a detecção e presença de uma aeronave intrusa, a uma distância de cerca de 1.3 km. [1]

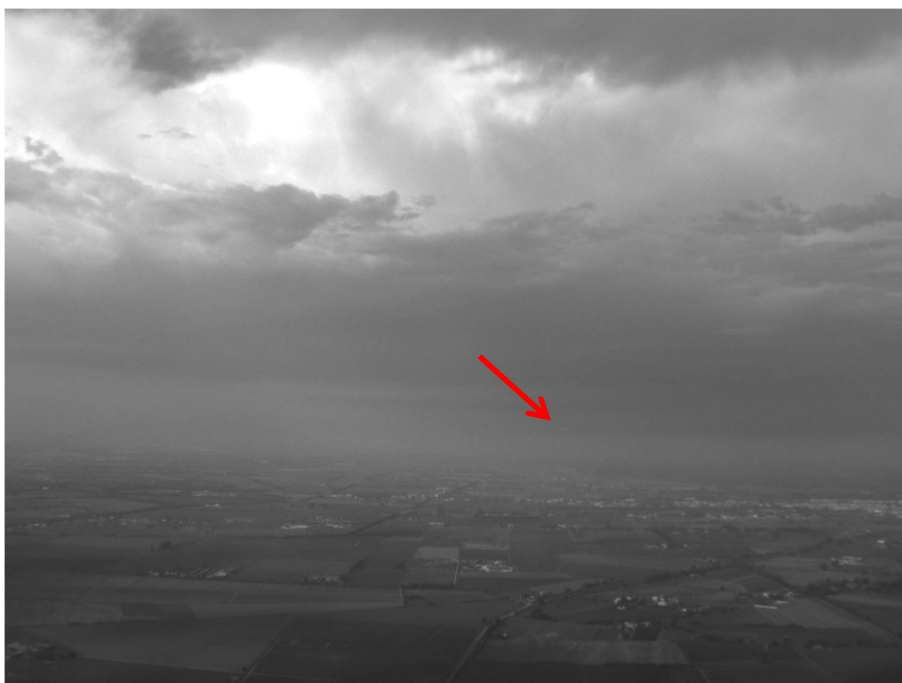


Figura 2.3: Imagem de um caso de teste com piores condições atmosféricas e de visibilidade do que as representadas na figura anterior. A seta vermelha indica a detecção e presença de uma aeronave intrusa, a uma distância de cerca de 1.3 km [1].

Tal como no artigo anterior, os autores de [12] também recorrem a um filtro de partículas para estimar a posição dos objetos detetados, sendo que, neste caso, o seu foco principal é lidar com objetos parcialmente oclusos. Os autores apresentam um método de seguimento de duas fases, baseado em análise de imagem, em cenários de elevada densidade visual, através de uma única câmara. Com o objetivo de encontrar a trajetória de cada objeto, assim como a trajetória global de todos os objetos, são iniciados dois estados: o estado local, e o estado global. No estado local é usado o filtro de partículas em conjunto com observação, para lidar com objetos parcialmente oclusos, que foram usados para gerar *tracklets* fiáveis. No estado global, são obtidos potenciais *tracklets*. Quando estes *tracklets* são relacionados pelo algoritmo Húngaro, obtém-se uma associação ótima global, resolvendo os problemas de associação de dados.

2.3 Seguimento de objetos com recurso a tecnologias ativas

Apesar das enormes vantagens e capacidades, que advêm do uso dos sensores passivos, nomeadamente, as câmaras, ainda subsistem muitas dificuldades em lidar com efeitos de sombras e reflexos, como é o caso de ambientes de superfície aquática. Uma outra desvantagem, que levou ao crescimento do desenvolvimento de tecnologias ativas, reside na dificuldade em medir, com alguma precisão, a distância à qual os objetos se encontram. Como forma de assegurar esta característica fundamental a um seguimento robusto em 3D, têm sido usadas tecnologias ativas para efeitos de mapeamento e de deteção.

Apesar da existência de métodos de seguimento, que utilizam sensores ativos para deteção de objetos, como, por exemplo, o RADAR e o SONAR, os resultados ainda não são bons, quando se pretende um sistema de multi-seguimento, devido à sua baixa resolução. Pelo contrário, o uso de sistemas baseados em laser, como é o caso da tecnologia LIDAR, apresentam bons resultados quando se procura a diferenciação e medição de objetos.

Um exemplo de seguimento, recorrendo a tecnologia LIDAR, é proposto em [2], onde é usado um sensor laser 2D para detetar, classificar e realizar seguimento de múltiplos objetos. Através da segmentação da nuvem de pontos, é possível identificar e classificar os diferentes tipos de objetos na estrada, como pode ser visto na figura 2.4.

Após a filtragem, classificação e determinação das posições dos diversos objetos no mundo, é então aplicado um filtro de Kalman, de forma a melhorar o processo de estimação dos veículos detetados, nomeadamente a posição e velocidade. Este sistema foi testado e validado com diversos sensores em situações reais, apresentando excelentes resultados, inclusive em situações de oclusão dos alvos.

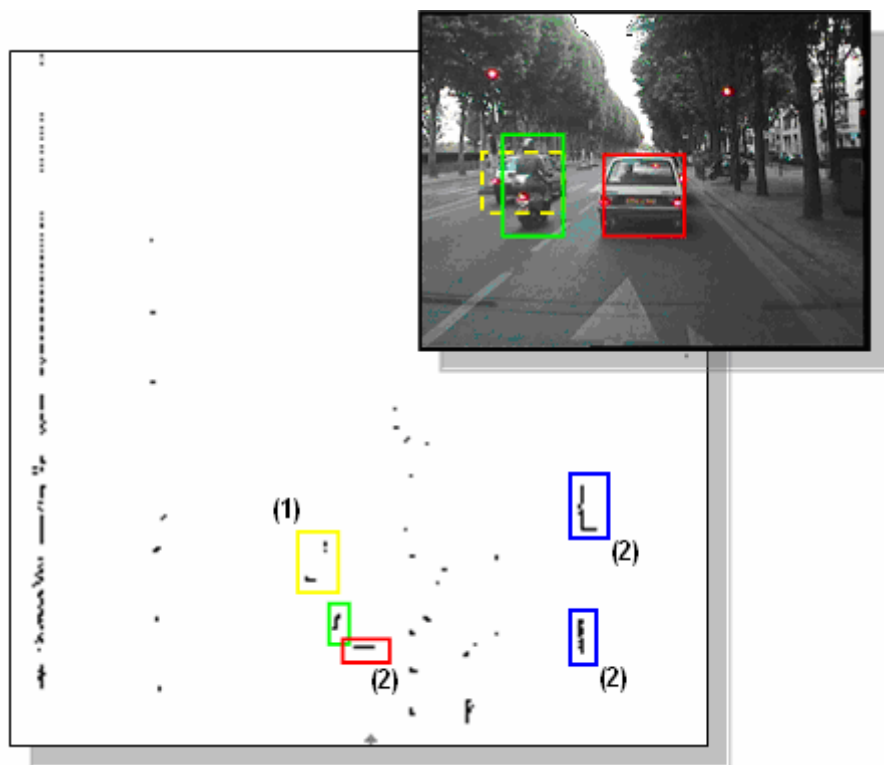


Figura 2.4: Detecção por LIDAR e a correspondente vista na imagem. Em 1: Carro ocluso. Em 2: Diversas formas de carros [2].

De forma semelhante ao artigo anterior, em [3], é apresentado um método de seguimento baseado exclusivamente em LIDAR. Neste sistema, é usado um sensor avançado, capaz de medir e registrar, com elevada definição, sequências de nuvens de pontos 3D a frequências de vídeo. Graças a essa capacidade, é possível ter uma elevada taxa de atualização de dados, concedendo ao sistema uma maior precisão. De forma a se obterem os candidatos para o processo de seguimento, submetem-se os dados anteriores a uma série de algoritmos, entre eles, o RANSAC, que eliminam os pontos pertencentes ao plano do solo. Desta forma, obtêm-se apenas os pontos que possuem uma maior probabilidade de serem considerados candidatos. Por último, são eliminados os pontos singulares, considerados como ruído, ficando apenas os candidatos claramente diferenciados, como é possível observar na figura 2.5. Uma vez adquiridos os centróides dos candidatos, através de um simples algoritmo de *clustering*, é aplicado um sistema de multi-seguimento baseado em filtros de Kalman, capaz de estimar e prever os movimentos e estados dos candidatos ao longo o tempo.

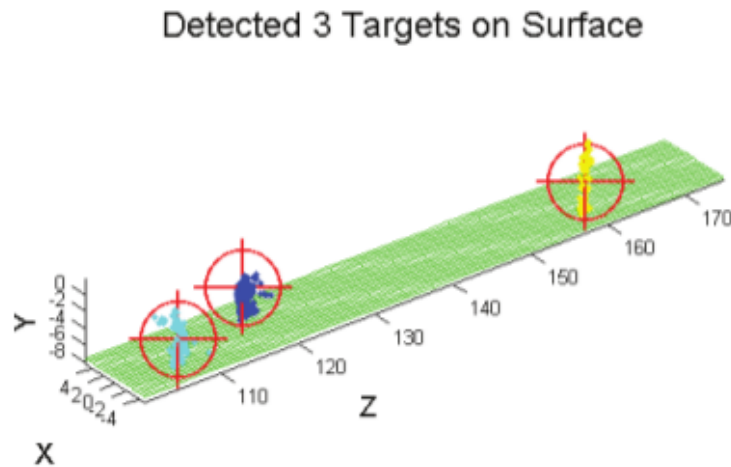


Figura 2.5: Através do algoritmo de *clustering*, foram determinados três alvos distintos[3].

2.4 Seguimento de objetos com recurso a tecnologias passivas e ativas

Com o intuito de juntar as vantagens e contrariar as desvantagens que advêm do uso isolado das tecnologias acima referidas, foram desenvolvidos sistemas multissensoriais. Estes sistemas incorporam tanto tecnologias passivas como ativas, numa tentativa de produzir um sistema mais robusto e preciso. Um exemplo deste sistema, é proposto em [4], onde um sistema de deteção e seguimento de obstáculos, para aplicações de segurança rodoviária. Este sistema recorre à fusão de dados provenientes de um sensor 3D LIDAR e de imagens 2D, provenientes de uma câmara. Apesar das deteções baseadas em LIDAR serem bastante precisas, não são, em termos de processamento, tão rápidas quanto as deteções efetuadas com base em imagens. Desta realidade, o método proposto usa a informação geométrica retirada da deteção por imagem, para extrair informações de profundidade, que são depois corrigidas pelas informações do LIDAR.

Após a fusão destes dados, e em ordem a realizar-se o seguimento, evitando os obstáculos presentes na estrada, é usado filtro de Kalman estendido, descrito em [13], onde um modelo de aceleração como ruído é considerado. O funcionamento deste sistema pode ser observado em 2.6.

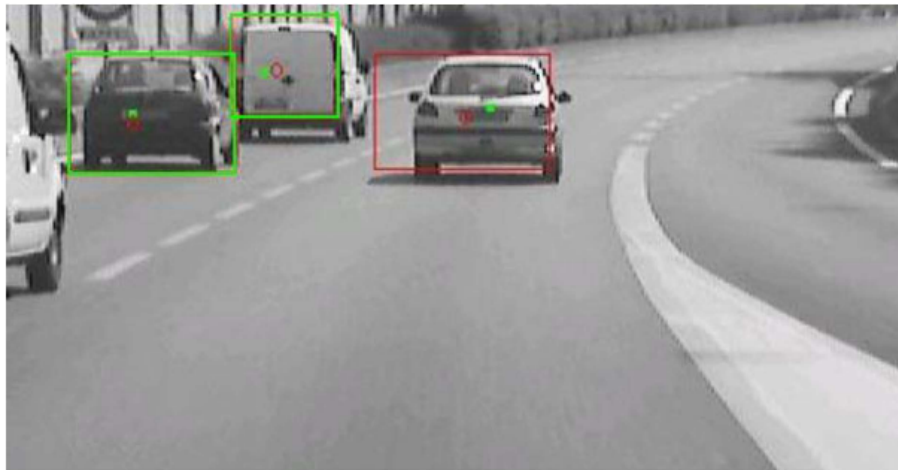


Figura 2.6: Detecção e seguimento de múltiplos obstáculos. A verde estão representadas potenciais ameaças, enquanto que a vermelho é expressa a ameaça de colisão mais iminente [4].

2.5 Conclusão

Após uma análise do uso das diversas tecnologias usadas na realização de seguimento de objetos, é possível concluir que o uso singular de cada tecnologia permite a criação de sistemas funcionais. No entanto, estes sistemas apresentam normalmente pouca robustez quando aplicados às diversas situações. Isto é, tanto o uso de câmaras como de *range finders* têm uma extensão de cenários de aplicação reduzida, uma vez que ainda apresentam grandes desvantagens que condicionam a operabilidade.

Tendo este fator em consideração, conclui-se ainda que o uso de tecnologias multissensoriais permite a criação de sistemas mais robustos e complementares de elevada precisão, uma vez que é possível, muitas vezes, contornar, ou reduzir substancialmente as desvantagens individuais de cada tecnologia. Assim sendo, o desenvolvimento dos sistemas de deteção e seguimento de objetos passará pelo uso e implementação de tecnologias multissensoriais.

Capítulo 3

Seguimento Multissensorial

Neste capítulo são apresentadas as tecnologias (*hardware* e *software*) utilizadas no processo de seguimento de objetos dinâmicos, assim como a descrição detalhada do algoritmo que permite o seu funcionamento. Após a exposição destes temas, serão ainda apresentados e discutidos os resultados do sistema desenvolvido.

Assim sendo, este capítulo encontra-se dividido nas seguintes temáticas:

- Arquitetura sensorial;
- Algoritmo de seguimento; e
- Resultados.

3.1 Introdução

O trabalho presente nesta dissertação tem como objetivo a criação de um sistema capaz de seguir e prever o movimento de objetos. A fim de atingir este objetivos, recorreu-se à fusão de dados de dois sensores, nomeadamente imagens provenientes de uma câmara e de uma nuvem de pontos do ambiente circundante fornecidos por um sensor LIDAR 3D.

É essencial ao funcionamento de qualquer método de seguimento a existência de uma entidade responsável pela deteção e identificação dos objetos, sobre os quais o sistema atuará. Não sendo a deteção e identificação de objetos o foco desta dissertação, foi utilizado um método baseado em processamento de imagem, explicado com mais detalhe na secção seguinte, para identificar e recolher dados de posicionamento de potenciais objetos alvo. Ao aliar a quantidade de dados transmitidos pela imagem da câmara com a precisão dos lados do laser é possível obter um posicionamento mais exato dos objetos em questão. Uma vez adquiridas as localizações, no mundo tridimensional, estas são alimentadas ao sistema para que este possa estimar e fazer uma gestão do movimento dos alvos ao seu redor.

De forma muito genérica é assim descrito o modelo de seguimento proposto nesta dissertação que será detalhado no seguimento da leitura deste capítulo.

3.2 Arquitetura sensorial

O método da obtenção e seguimento de objetos foi desenvolvido com recurso a diversas ferramentas, desde bibliotecas a um metassistema operativo direccionado para aplicações robóticas. Este metassistema operativo, o ROS ¹, foi usado como meio de interligação entre os diversos componentes do sistema. A razão da sua seleção deveu-se ao facto de o módulo de aquisição de dados, no qual estão incluídos a câmara e o LIDAR, já se encontrar, também ele, a funcionar sobre este ambiente.

O ROS, graças ao seu vasto conjunto de bibliotecas *open-source* e ferramentas, possibilita a construção das mais variadas aplicações robóticas, uma vez que este facilita a gestão dos processos de integração e comunicação dos vários módulos do sistema. Na figura 3.1, é exposto uma vista geral da interligação do sistema de perceção usado.

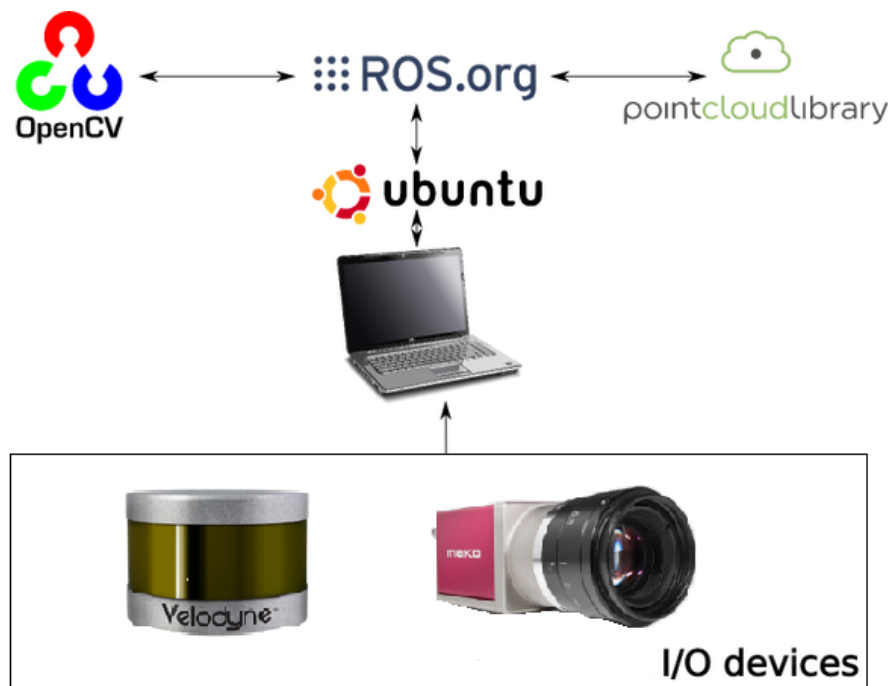


Figura 3.1: Vista geral do sistema e das ferramentas utilizadas

Esta interligação entre os diferentes constituintes é alcançada através da criação de nós que permitem a entrada de informação, realizam o seu processamento e originam determinados *outputs* [14]. Esta rede de gestão de informação é baseada na arquitetura de *software* publicador - subscritor. Uma ilustração deste tipo de interação é visível na figura 3.2, onde é possível identificar os três principais componentes de uma interação ROS, nomeadamente:

- **Nós** - Programas independentes capazes de publicar e subscrever mensagens em tópicos;

¹<http://www.ros.org/>

- **Mensagens** - Entidade individual sobre a qual a comunicação entre nós é feita; e
- **Tópicos** - Canal de transmissão usado pelos diversos nós.

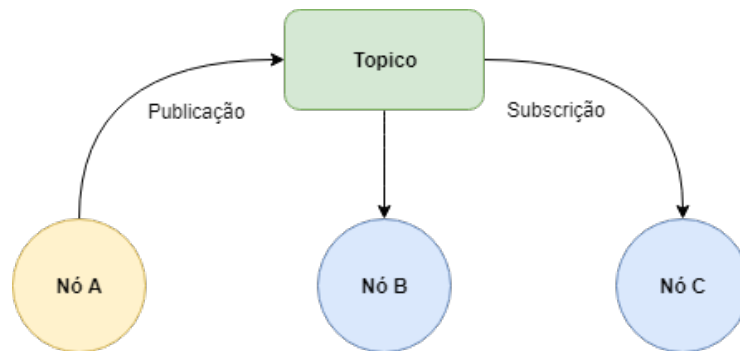


Figura 3.2: Exemplo ilustrativo de uma interação entre 3 nós, no ROS

O uso deste tipo de estrutura de ligação, entre os diversos componentes de um sistema, oferece uma outra vantagem - a modularidade. Ao permitir a modularidade dos diferentes intervenientes, todo o processo de melhoria, adaptação e substituição é facilitado. Desta forma, a integração de outras bibliotecas como, nesta situação em particular, a PCL e a OpenCV, é facilmente alcançado, munindo o sistema de uma grande versatilidade e adaptabilidade nos termos do uso de ferramentas e câmbios de dados.

Dando agora uma maior atenção às bibliotecas usadas, a biblioteca do OpenCV é usada para aplicações de visão, sendo utilizada nas mais diversas aplicações desta área. Esta permite, entre muitas outras, a filtragem de imagens, a calibração de câmaras e o reconhecimento de objetos. O OpenCV é altamente compatível e oferece suporte para diversos sistemas operativos, estando disponível em linguagens como C++, Python e Java. Os seus algoritmos podem ser usados por terceiros a nível integral e parcial, sendo que a sua essência assenta na aplicação de sistemas de tempo-real e eficiência de recursos computacionais.

O uso da biblioteca OpenCV na presente dissertação, relaciona-se com a necessidade de processar as imagens transmitidas pela câmara, afim de se conseguir uma identificação dos objetos e uma estimativa primária da sua posição. Como foi dito anteriormente, a identificação e o reconhecimento de objetos são técnicas utilizadas pelo OpenCV, sendo por isso a primeira escolha, quando se trata de extração de dados de imagens.

Por fim, resta apenas referir a existência de uma outra biblioteca, que em integração com o ROS completa o sistema final - a PCL [15]. Esta biblioteca permite a manipulação de nuvens de pontos de elevadas dimensões. Uma nuvem de pontos é um conjunto de pontos representado num determinado sistema de coordenadas. Sendo que o sistema desenvolvido é tridimensional, estes pontos são definidos nas coordenadas x, y e z. Com o auxílio desta biblioteca, é possível caracterizar as superfícies externas dos objetos presentes no ambiente circundante, adquirindo informações de elevada exatidão do posicionamento dos objetos. Assim sendo, a PCL tem um papel fundamental no processamento das nuvens de pontos recebidas pelo sistema de aquisição.

3.2.1 Sensores do Sistema de Aquisição

Como se pode ver na figura 3.1, os sensores utilizados pelo sistema, nomeadamente para a extração dos dados do exterior são: uma câmara e um LIDAR.

A câmara utilizada corresponde a uma Mako G-125 (figura 3.3) e as suas especificações encontram-se descritas na tabela 3.1. Esta câmara, construída para aplicações robóticas é capaz de fornecer, em tempo real, imagens monocromáticas e policromáticas, através de uma conexão *ethernet* RJ45.

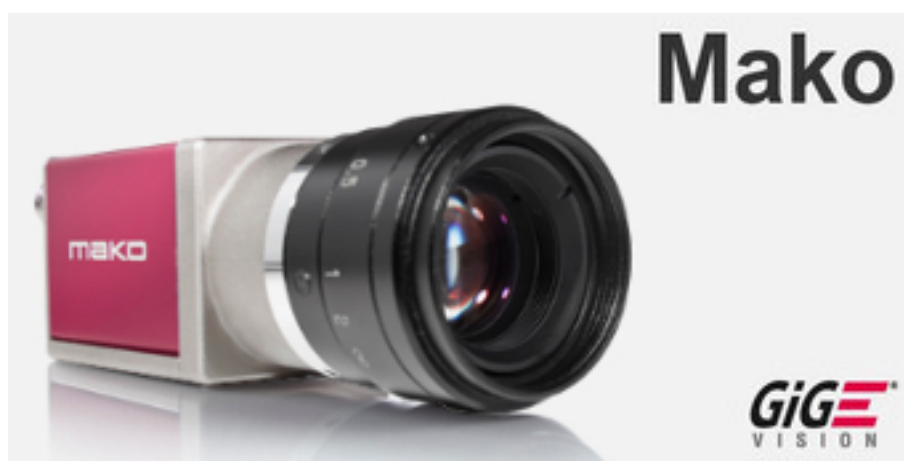


Figura 3.3: Câmara - Mako G-125 [5]

Tabela 3.1: Especificações do Mako G-125 [5]

Resolução	1292 (H) x 964 (V)
Tipo de Sensor	CCD Progressivo
Tamanho do sensor	1/3
FPS max	30.3 fps
Peso	80 g

O sensor LIDAR mencionado corresponde ao sensor VLP-16 (figura 3.4). O VLP-16 é um sensor capaz de fornecer dados tridimensionais em tempo real. Com um alcance máximo de 100 metros e um campo de visão de 360°, este sensor adequa-se perfeitamente à tarefa ao qual é aplicado. Alguns dados mais específicos encontram-se explícitos na tabela 3.2.

A escolha destes sensores, para integração do sistema de seguimento, prendeu-se com o facto de que estes já se encontrarem montados e calibrados no laboratório de investigação do CRAS do INESC TEC, no polo da FEUP.



Figura 3.4: Sensor LIDAR - VLP-16 [6]

Tabela 3.2: Especificações do VLP-16 [6]

Canais	16
Alcance	100 m
Precisão	+/- 3 cm
FOV vertical	+15° até -15° (30°)
Resolução angular vertical	Vertical: 2°
Resolução angular horizontal	Horizontal: 0.1° - 0.4°
FOV Horizontal	360°
Número de Pontos gerados	300.000 - 600.000 p/s

3.3 Algoritmo de Seguimento

3.3.1 Introdução

Apresentadas as tecnologias usadas para o desenvolvimento e implementação do sistema, resta agora, com mais detalhe, explicar todo o processo desenvolvido, assim como o enquadramento deste projeto na área da aplicação robótica. Começando por este último ponto, esta dissertação pretende desenvolver um sistema de seguimento baseado em fusão multissensorial, com o objetivo de vir a ser integrado num ASV. Este, com a integração do sistema desenvolvido, terá a capacidade de detetar diversos objetos, podendo estimar as suas trajetórias e proceder, se assim o desejar, à perseguição de um deles. Podendo um ASV ser um sistema constituído por múltiplos subsistemas, é importante referir que o trabalho desenvolvido é totalmente focado no módulo de seguimento. Assim sendo, o sistema terá como tarefa processar um dado número de pontos 2D (x,y) , fornecidos pelo módulo de deteção e apresentar na saída um conjunto de pontos 3D (x,y,z) . Este último conjunto de pontos representa as previsões das trajetórias dos objetos seguidos, sendo estes fornecidos ao módulo de navegação. Uma ilustração generalista da descrição prévia pode ser visualizada na figura 3.5, na qual é retratado e enquadrado o tema desta dissertação, assim como os seus dados de entrada. Por sua vez, na figura 3.6, é apresentado, sobre a forma de fluxograma, todas as etapas do algoritmo desenvolvido, que será explicado com mais detalhe, ao longo desta secção.

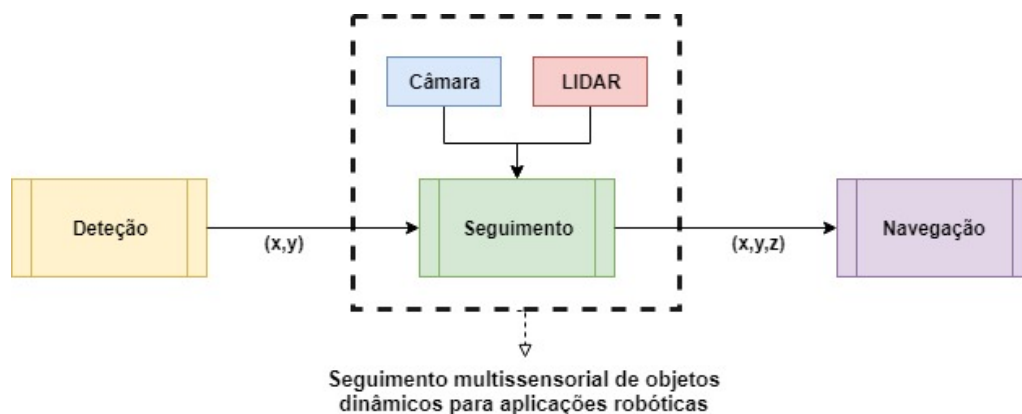


Figura 3.5: Enquadramento da dissertação

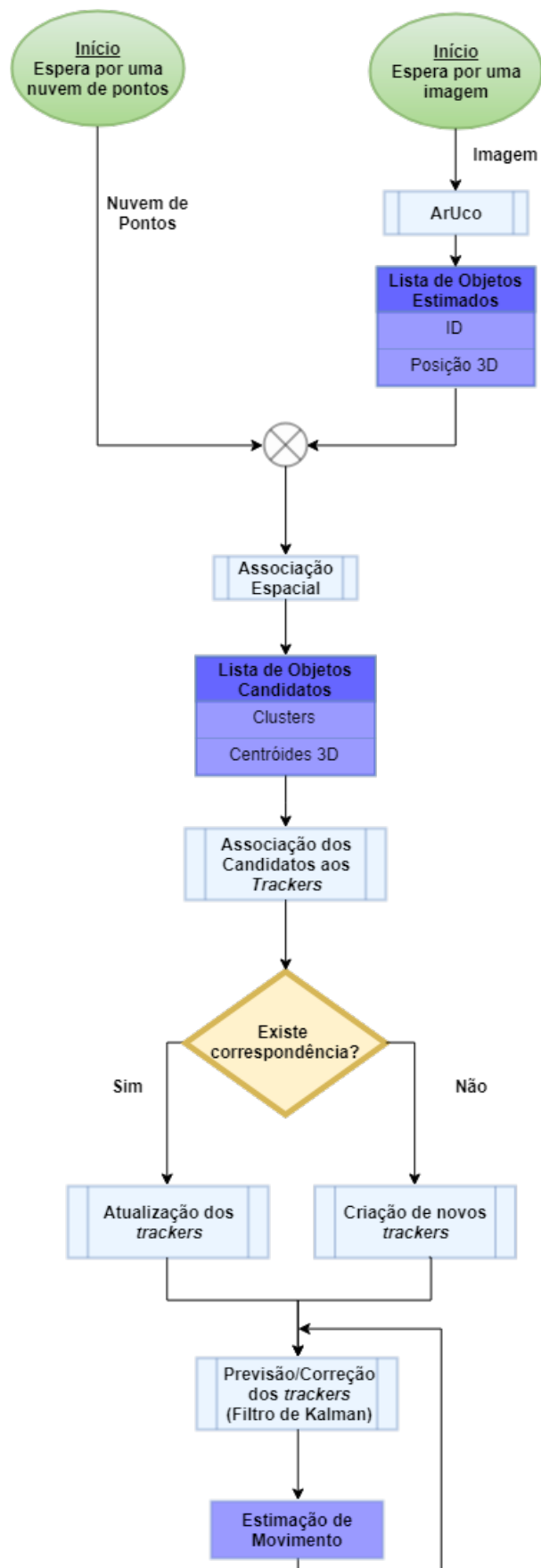


Figura 3.6: Fluxograma das etapas do algoritmo desenvolvido

3.3.2 Metodologia e Especificações Técnicas

Apesar de o foco deste desenvolvimento ser apenas o algoritmo de seguimento, é necessário um sistema de deteção capaz de o alimentar. Desta forma, foi necessário a procura de um módulo de deteção capaz de satisfazer o maior número de requisitos. Esses requisitos pretendidos traduzem-se na capacidade de detetar múltiplos objetos e retornar as suas posições, na necessidade de baixos recursos computacionais para efetuar a deteção, ser facilmente integrável com o ROS e, acima de tudo, ser capaz de efetuar a deteção e processamento de dados em tempo-real.

Após a análise dos requisitos descritos e dos sistemas existentes, a escolha recaiu sobre uma deteção baseada em imagem, sendo escolhido o ArUco [16, 17] como sistema de deteção primária. O ArUco é uma biblioteca *OpenSource*, integrada no OpenCV, que deteta a presença de marcadores, visíveis na figura 3.7, em imagens. Um marcador ArUco, visível na figura 3.8, é um marcador quadrado composto por uma borda preta com uma matriz binária no seu interior que permite determinar o seu identificador (ID). O processo de deteção deste tipo de marcadores pode ser dividido em dois passos. O primeiro consiste na procura do candidato, onde são procuradas, na imagem, formas semelhantes a quadrados. A borda preta facilita uma deteção rápida na imagem. A codificação binária permite a sua identificação e a aplicação de técnicas de deteção e correção. O tamanho de um marcador é determinado pelo tamanho da matriz interna. Por exemplo, um marcador com tamanho 4x4 bits é composto por uma matriz quadrada de 16 bits. O segundo passo no processo de deteção consiste na descodificação do marcador em si. O marcador candidato é analisado a fim de se verificar se este é de facto um marcador e, em caso positivo, se é um marcador de determinado conjunto de marcadores, também conhecido como dicionário. Um dicionário não é mais do que uma lista que contém as codificações para cada marcador.

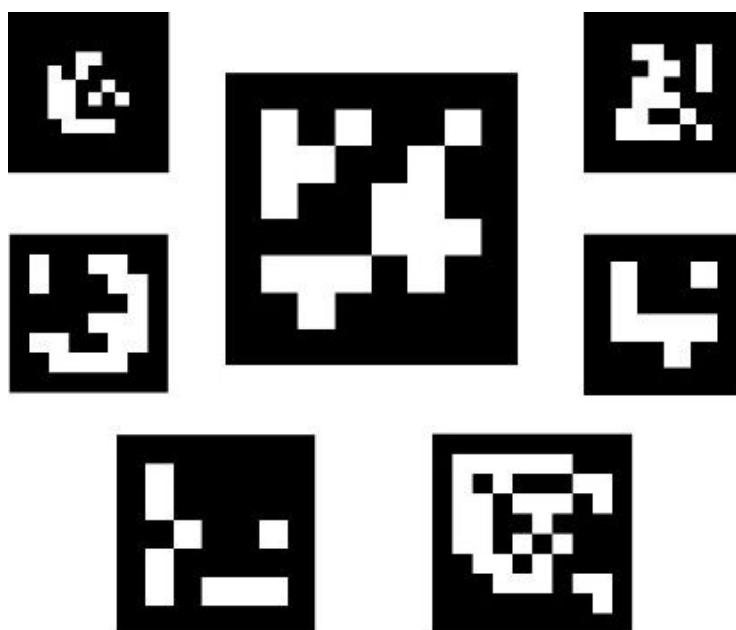


Figura 3.7: Exemplos de marcadores [7]

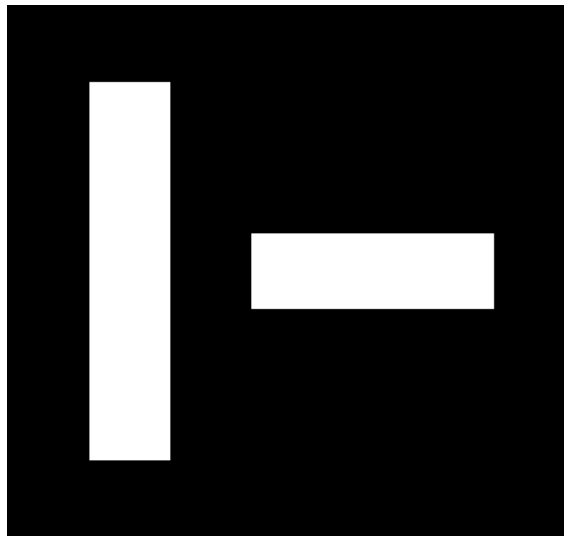


Figura 3.8: Marcador ArUco:
ID - 16; Dicionário - 16; Dimensão - 6x6

Após a deteção dos marcadores, o ArUco, através dos parâmetros de calibração da câmara, como a matriz dos parâmetros intrínsecos da câmara e os coeficientes de distorção, é capaz de obter a pose câmara em relação aos marcadores. Assim é possível gerar, para cada marcador, a posição, em coordenadas (x,y,z) , e a orientação do seu centróide na imagem. É de notar que este sistema é capaz de detetar os marcadores independentemente da rotação que estes apresentam. Na figura 3.9, é possível observar o *output* do sistema na deteção de dois marcadores, identificando-os pelo ID, e mostrando a sua orientação e rotação através de um sistema de eixos.

A escolha do ArUco, como sistema de deteção, deveu-se ao facto de este ser um sistema capaz de cumprir todos os requisitos estabelecidos e fornecer, para além da posição do objeto, um identificador do mesmo, facilitando o processo de *debugging*.

Assim sendo, existe agora um sistema capaz de analisar o meio exterior e fornecer ao módulo de seguimento, as posições tridimensionais dos objetos em relação ao sistema. No entanto, como o cálculo da posição 3D do objeto é feito através do processamento de uma imagem 2D, existe um erro associado ao cálculo da profundidade. O cálculo da distância do marcador à câmara é realizado, através da comparação da medida do comprimento do marcador com o seu tamanho real, que é conhecido e igual entre marcadores do mesmo dicionário. Este erro é bastante reduzido quando se usa o marcador impresso no seu tamanho padrão. Todavia, este uso implica uma distância máxima de operação de aproximadamente dois metros. Este teto máximo de operabilidade deve-se ao facto de o sistema não conseguir descodificar o padrão binário, uma vez que a imagem fornecida pela câmara não possui definição suficiente. Desta forma, existe a necessidade de aumentar o tamanho dos marcadores para que estes possam ser detetados a distâncias maiores, uma vez que a distância máxima anterior não é compatível em aplicações de navegação autónoma, tanto em termos de operação como de segurança. Consequentemente, o aumento do tamanho do marcador, para além do tamanho padrão, origina um aumento substancial do erro provocado, que

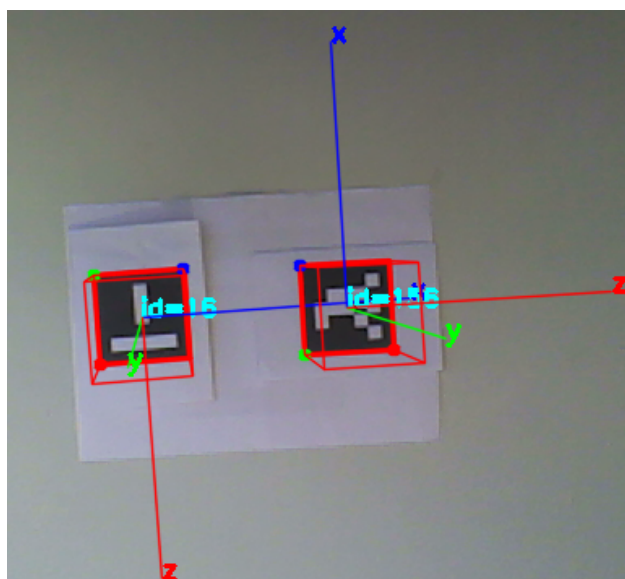


Figura 3.9: Exemplo de interação do Aruco com dois marcadores

é tanto maior quanto o aumento das dimensões do marcador. Com este erro, o sistema indicará que os objetos estão mais próximos da origem do sistema do que na realidade estão. Por conseguinte, o sistema deixa de poder fornecer posições tridimensionais precisas, na ordem dos centímetros, dos objetos detetados, proporcionando apenas uma zona aproximada de onde o objeto poderá estar localizado. Esta situação não é compatível com um seguimento adequado, uma vez que a estimação da trajetória do objeto apresentaria um erro demasiado elevado, condicionando severamente o correto funcionamento do sistema que dele dependesse.

Com a necessidade de obter dados fiáveis com os quais se possa realizar um seguimento, recorreu-se à fusão dos dados provenientes ArUco com os dados provenientes de uma nuvem de pontos do ambiente circundante. Usando a tecnologia LIDAR, nomeadamente o sensor VLP-16, descrito em 3.2.1, é possível obter uma mapa 3D detalhado do ambiente circundante do sistema, como é possível observar na figura 3.10.

Com um mapa tridimensional detalhado do ambiente em redor e uma posição aproximada dos objetos alvos, é possível a extração exata do centróide de cada objeto. Para tal, foi necessário primeiro transferir todos os valores do sistema de deteção e da nuvem de pontos para o mesmo referencial. Este passo é crucial, uma vez que, como os dados provêm de fontes diferentes, têm naturalmente referenciais diferentes, o que impede a sua associação. Para realizar esta tarefa, recorreu-se a uma ferramenta bastante conhecida do ROS - a *tf* [18]. A *Transform Library* ou *tf* é uma biblioteca que permite manter um controlo sobre os referenciais e transformar os dados existentes de um sistema de tal forma que qualquer membro do sistema poderá usar esses dados dentro do seu próprio referencial sem necessitar de conhecer todos os referenciais existentes. Para realizar esta transformação, foram utilizados os parâmetros extrínsecos da câmara e do LIDAR em relação a um referencial que se considera como a origem física do sistema. Estes parâmetros tra-

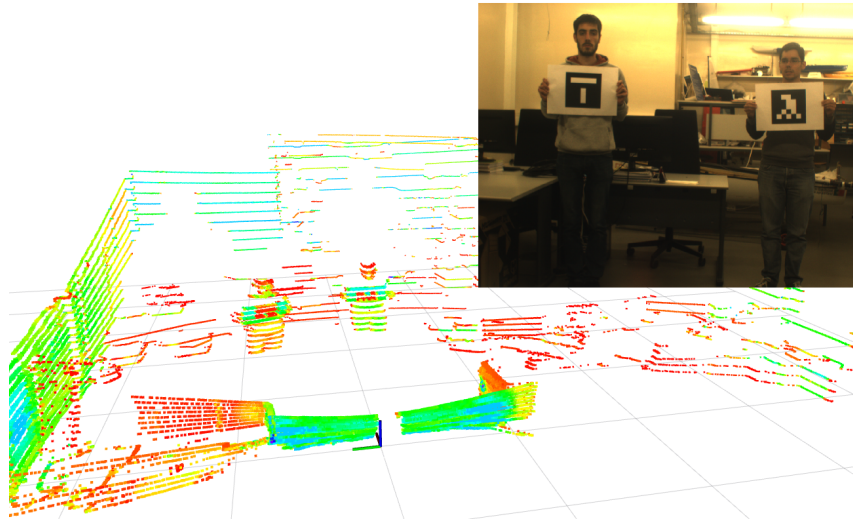


Figura 3.10: Nuvem de Pontos do ambiente

duzem a posição e orientação entre os referenciais e, assim sendo, os novos pontos são o resultado da expressão matemática apresentada em 3.1. Nesta equação, a variável R representa a matriz de rotação. $[X, Y, Z]$, as coordenadas do ponto original e T constitui a matriz de translação. Estas informações foram geradas aquando da montagem e calibração do sistema, sendo disponibilizadas via ROS.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T \quad (3.1)$$

Com a aplicação desta ferramenta, consegue-se garantir que um ponto gerado pela câmara é identificável na nuvem de pontos, concluindo-se, assim, a primeira etapa da construção do sistema. O seu estado atual é observável na figura 3.11.

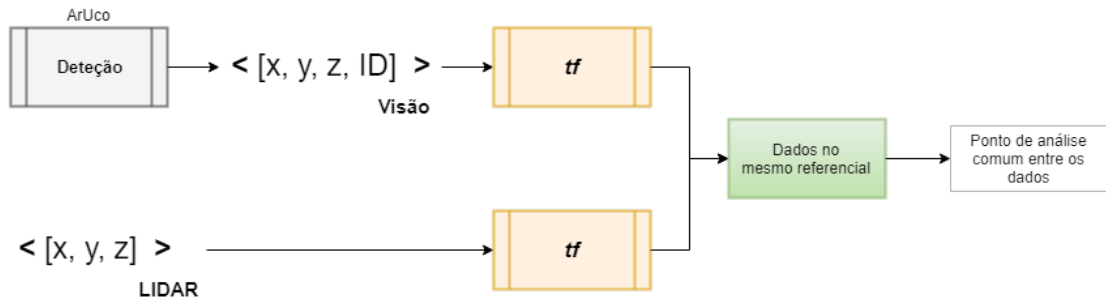


Figura 3.11: Transformação e fusão dos dados sensoriais para um referencial comum

Uma vez realizada a transformação dos dados para os referencial comum, é possível agora proceder ao processamento e fusão dos mesmos. Este processo traduz-se na identificação, na nuvem de pontos, do objeto alvo detetado pela câmara, através do uso de funções de processamento de nuvens de pontos, disponíveis na biblioteca PCL.

Uma vez que, devido ao erro gerado no sistema de deteção, a posição estimada do marcador não corresponde à verdadeira posição do objeto na nuvem de pontos. Esta discrepância é visível na figura 3.12, onde se pode observar, para cada objeto, a diferença entre a posição estimada pelo ArUco, representado pelas esferas amarelas, e a posição real dos objetos no mundo, representado pelos paralelepípedo a preto. Contudo, é sabido, e observável que o erro só se relaciona com a distância do objeto à origem do referencial. Desta forma, é possível extrapolar uma região onde o objeto se encontra e determinar a sua posição.

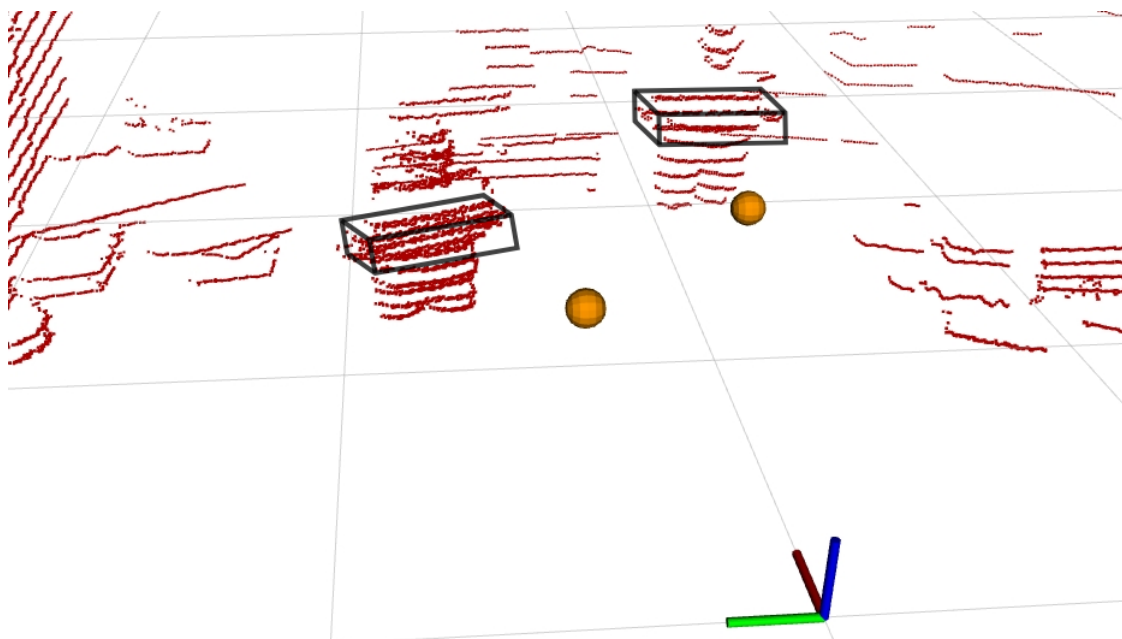


Figura 3.12: Discrepância sensorial:
Esferas Amarelas - Posições estimadas;
Paralelepípedo Pretos - Posições reais.

O processo determinação da posição do centróide de cada objeto pode ser dividido em quatro partes distintas:

- **Segmentação da nuvem de pontos num ROI 3D**

Após a experimentação e análise do ArUco com os marcadores alterados, verificou-se, que este estima, em média, a posição do objeto com um erro aproximado de 30%. Isto significa que a distância real à origem do referencial a que o objeto se encontra é aproximadamente 1.3 vezes superior à distância apontada. Tendo em conta esta variável foi aplicado um processo de filtragem à nuvem de pontos. Este processo consiste num filtro passa banda, segmentando a nuvem de pontos num determinado intervalo de coordenadas x, y ou z. Assim, com a aplicação de três destes filtros, um para cada eixo, foi possível a criação de um ROI 3D, na nuvem de pontos, do objeto estimado. O resultado da aplicação desta filtragem pode ser analisado, através da comparação das figuras 3.13 e 3.14, onde se observa a segmentação da nuvem de pontos original (figura 3.13) e a criação de uma nova nuvem de pontos, da região de interesse (figura 3.14), onde se pode constatar a presença do objeto que se pretende seguir.

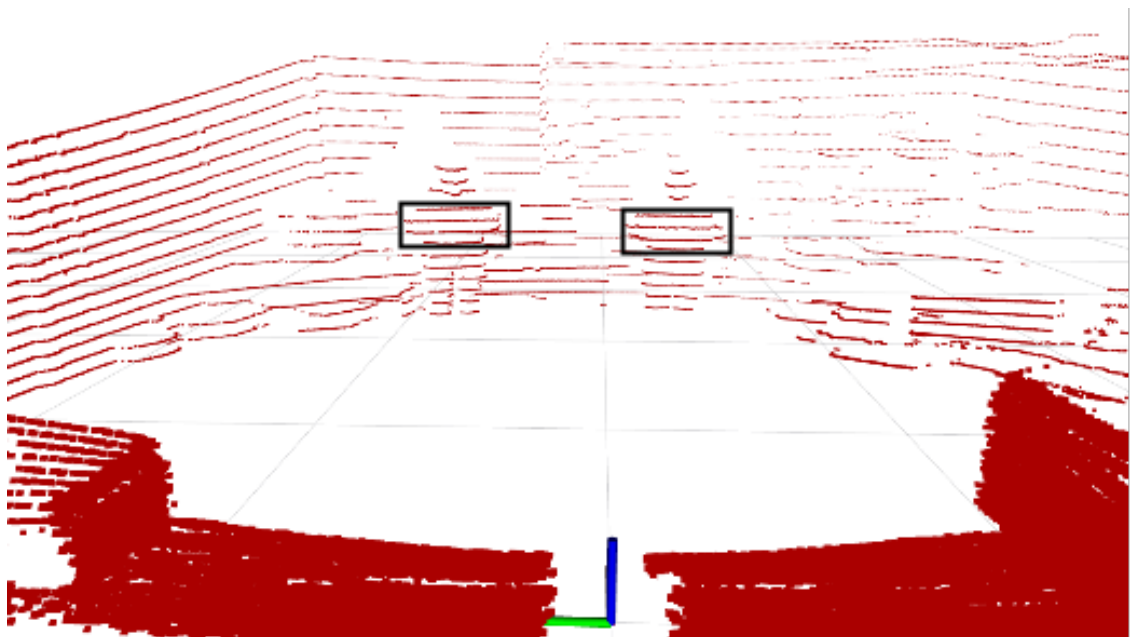


Figura 3.13: Nuvem de pontos antes da segmentação

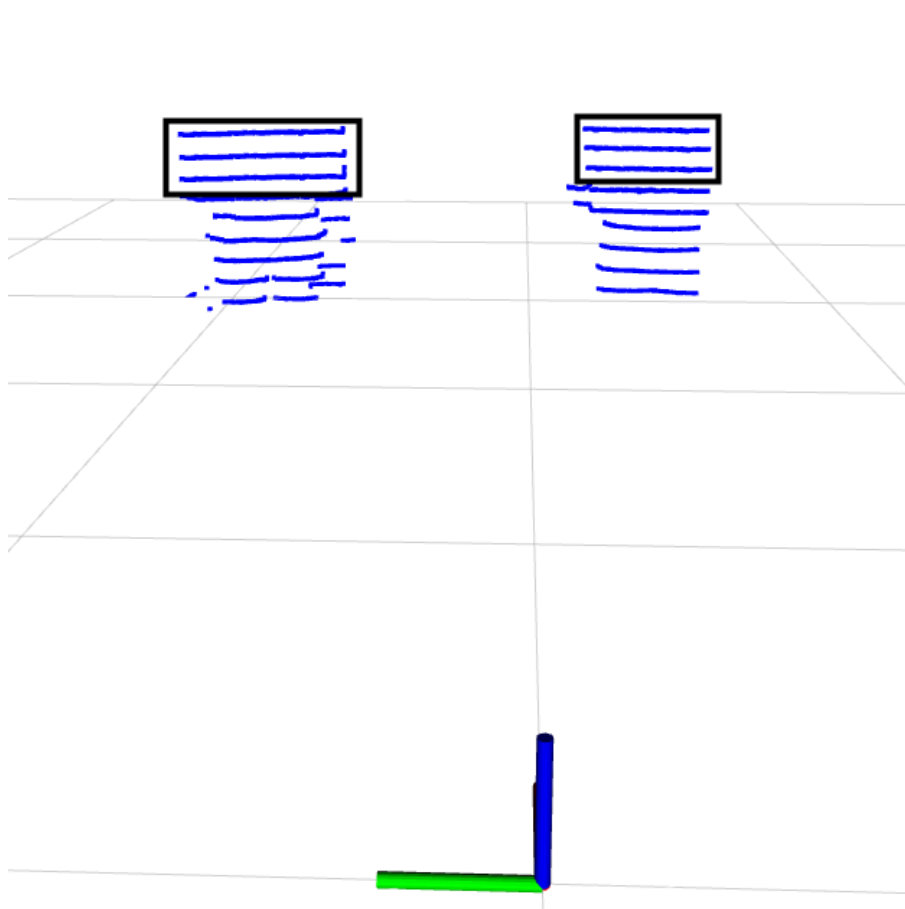


Figura 3.14: Região de interesse obtida após filtragem

- **Formação e extração de *clusters***

Após a criação de uma nuvem de pontos mais restrita, uma vez que contém um reduzido número de pontos, quando comparada com a nuvem de pontos original, é necessário organizá-la em aglomerados ou *clusters*. Esta organização da nuvem de pontos permite a diferenciação entre os diferentes corpos existentes sendo, neste caso, o objetivo diferenciar o objeto candidato dos restantes corpos circundantes.

O método de extração euclidiana de *clusters* foi o utilizado para se proceder à realização desta tarefa. A grande vantagem no uso deste método é a sua capacidade de dividir uma nuvem de pontos desorganizada em pequenas partes, reduzindo significativamente o tempo de processamento da mesma. Com a execução deste algoritmo de *clustering* obtém-se, para cada região de interesse, um ou mais *clusters*, assim como os pontos pertencentes a cada um deles.

- **Cálculo do centróide de cada *cluster***

Sendo um *cluster* um agregado de pontos que pode assumir qualquer forma ou tamanho, é necessário, para efeitos de processamento e representação, reduzir cada um deles a um único ponto, neste caso, o seu centróide. Para isso, recorreu-se ao cálculo da média aritmética de todos os pontos de cada *cluster*, obtendo-se, como pretendido, o centróide tridimensional de cada *cluster*.

- **Minimização dos centróides**

Quando é identificado mais que um *cluster*, torna-se imperativo a seleção de um, uma vez que apenas um deverá corresponder ao objeto candidato. O processo de seleção do candidato é realizado com recurso à minimização global das distâncias euclidianas entre o centróide de cada *cluster* e a posição tridimensional estimada pelo ArUco. Por outras palavras, considera-se como *cluster* do objeto candidato aquele que possui o centróide mais próximo da posição estimada pelo sistema de deteção.

Com a aplicação do método descrito nos pontos anteriores, é assim possível determinar, com um elevado grau de confiança, a posição tridimensional que o objeto ocupa no mundo em relação ao sistema. Na figura 3.15, é possível observar um fluxograma representativo dos diversos processos a que os dados são sujeitos, partindo dos dados fornecidos pelo sistema de deteção, nomeadamente o vetor de posições estimadas, até à obtenção do novo vetor de posições reais resultante da fusão sensorial. O resultado da aplicação deste método é visível na figura 3.16, onde se encontram representados tanto as posições estimadas pelo ArUco como também as posições reais dos objetos e os respetivos *clusters*, calculadas por este método.

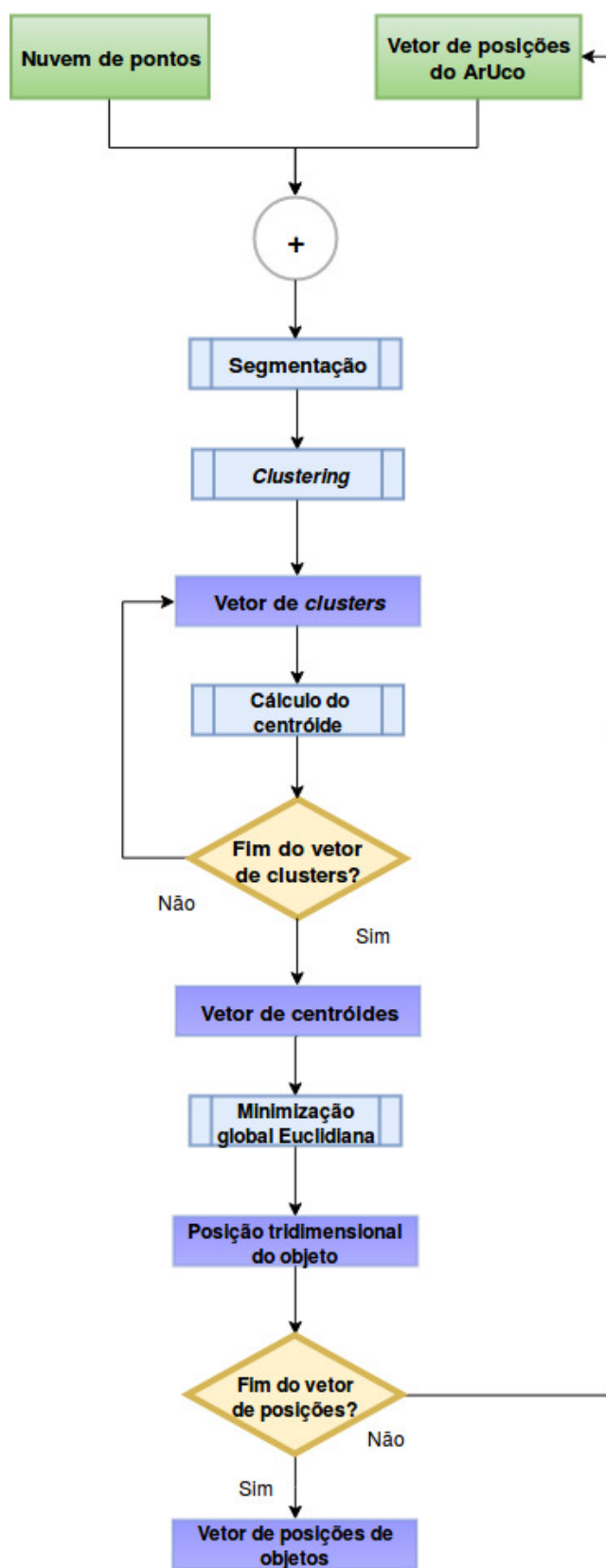


Figura 3.15: Algoritmo de fusão sensorial

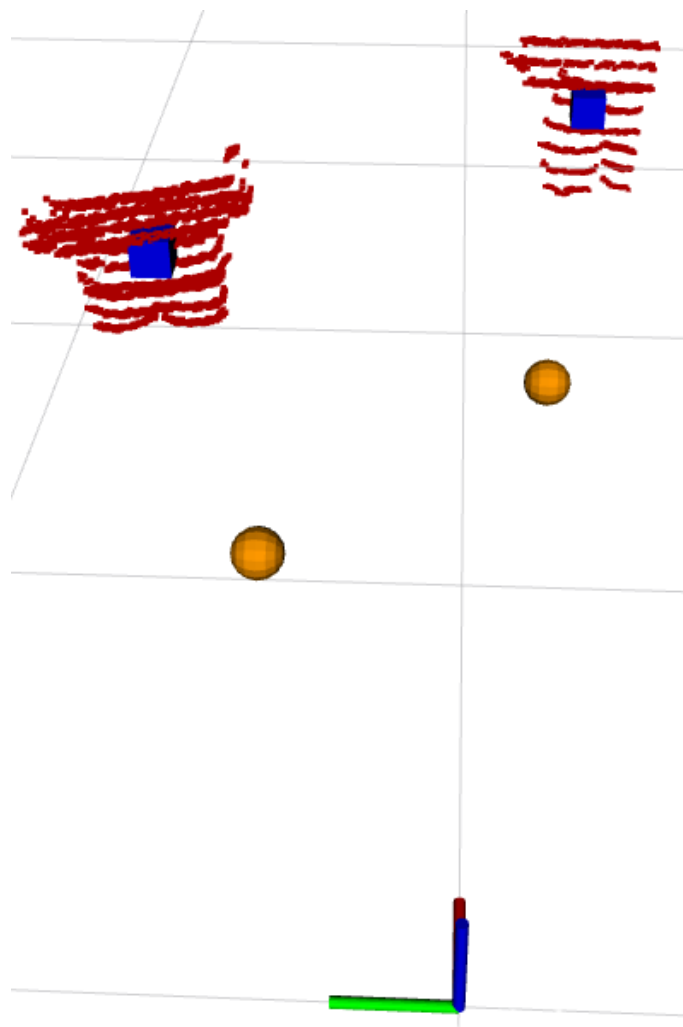


Figura 3.16: Resultado do algoritmo de fusão sensorial:
Esferas Amarelas - Posições estimadas;
Cubos Azuis - centróides computados.

Concluída a fase de determinação das posições tridimensionais dos objetos detetados, inicia-se agora a fase de seguimento dos mesmos. O método utilizado nesta dissertação para realizar o seguimento de objetos detetados, deriva do método apresentado em [8]. Os autores, para além de proporem uma arquitetura de um sistema robótico de vídeo-vigilância, apresentam ainda uma técnica de seguimento visual de múltiplos objetos. Tendo sido esta técnica direcionada ao seguimento de objetos detetados apenas por imagens, foi construída e otimizada para o processamento de pontos bidimensionais. Por esta razão, foi aplicado a este algoritmo um processo de conversão para que pudesse operar sobre os pontos tridimensionais fornecidos. Apesar das alterações realizadas, foi garantido que o núcleo do algoritmo se mantivesse intacto. Dessa forma, o método exposto em [8] mantém-se inalterável do ponto de vista algorítmico.

O método aplicado é chamada MTMP (Multi-Tracking of Motion Profiles) e encontra-se retratado na figura 3.17 o seu algoritmo lógico.

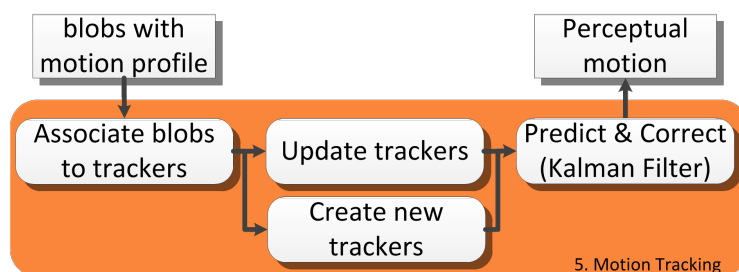


Figura 3.17: Diagrama lógico do algoritmo de seguimento multi-objeto [8]

Para uma correta compreensão do algoritmo apresentado, é importante a explicação de um conceito chave. Este conceito relaciona-se com a definição de "*blob*", uma vez que este representa uma peça fulcral na operabilidade, não apenas do método de seguimento mas do sistema como um todo. Um *blob*, não é mais do que a representação do centróide de um objeto tridimensional detetado, e que foi fornecido ao módulo de seguimento.

Este método é capaz de estimar o perfil de movimento de objetos ao recorrer a múltiplos Filtros de Kalman, que para além de conferirem a este sistema a capacidade de seguir vários alvos, confere ainda a capacidade de executar este processo com reduzidos recursos computacionais.

A característica chave deste método é a atribuição de um filtro de Kalman para cada *blob*. Esta associação forma uma nova entidade, o *tracker*, que se caracteriza pelos seus modelos de movimento de velocidade constante e por equações de medição das posições dos centróides/*blobs*. Como pode ser observado na figura 3.17, o processo encontra-se dividido em três fases distintas, sendo elas: associação, atualização dos *trackers* e previsão/correção. De maneira a se garantir o correto funcionamento deste sistema de seguimento, é necessário um complexo processo de gestão que permita gerir as relações entre os *blobs* e os vários filtros de Kalman. Para realizar esta tarefa recorre-se à criação de três tipos de *trackers*, nomeadamente: novos, ativos e inativos, sendo que cada um está associado a um, e um só *blob*, e consequentemente um perfil de movimento. A característica que permite a distinção entre estes tipos é a taxa de incidência ou observação de

um dado *blob*, sendo-lhe diretamente proporcional. Esta foi introduzida como forma de eliminar falsos positivos gerados pelos módulos anteriores, que poderiam causar uma excessiva criação de *trackers*, ou influenciar o cálculo do perfil de movimento dos *trackers* já existentes. Esta filtragem é alcançada pois só os *trackers* com uma taxa de incidência pré-estabelecida (*trackers* do tipo ativo) é que são consideradas para análise de perfil de movimento. Assim sendo, um *tracker* "Novo" é criado, sempre que o perfil de movimento de todos os *blobs* não correspondem ao de nenhum *tracker* existente. Um *tracker* já existente reforça o seu perfil de movimento sempre que deteta um *blob* com um perfil semelhante, nomeadamente a posição do seu centróide. Este reenforço do *tracker* causa naturalmente uma subida no valor da taxa de incidência. Desta forma, *trackers* com elevada taxa de incidência traduzem uma maior perceção do perfil de movimento, sendo considerados assim de - Ativos. De forma díspar, se a taxa de incidência do *tracker* descer abaixo de um certo limiar, este passa a constituir um *tracker* - Inativo, uma vez que traduz um possível desaparecimento do *blob*. No entanto, havendo ainda a possibilidade de este não aparecer por razões como a sua oclusão prolongada, é estabelecido um novo limite para a taxa, que se ultrapassado, causará a eliminação do *tracker*. É assim, graças ao uso do conceito de taxa de incidência que é possível realizar uma gestão eficiente e exata dos *trackers* existentes.

Entrando agora mais em detalhe, relativamente a cada um das três fases deste método, a da associação é a primeira, uma vez que é nela que são analisados os *blobs*. Estes são observados e comparados com os *trackers* existentes, procurando por aqueles que apresentam perfis de movimento semelhantes. Para realizar esta comparação é criada uma matriz de similaridade que contém as distâncias euclidianas entre todos os *blobs* e todos os *trackers*. A atribuição de um *blob* a um *tracker* depende se o valor da distância euclidiana entre eles é inferior a um valor pré-determinado, que se relaciona com o deslocamento máximo que se considera que um dado objeto pode realizar no tempo entre iterações. Desta forma, a operação de associação é realizada iterativamente tendo em conta as semelhanças entre perfis de movimentos e as distâncias euclidianas entre as posições dos centróides. Terminada a associação entre os *blobs* e os *trackers*, estes últimos são atualizados com os valores dos *blobs* ao qual foram correspondidos anteriormente. Este evento causará, como foi dito previamente, uma variação nos valores das taxas de incidência dos *trackers* assim como a possível criação e/ou eliminação de outros.

Uma vez atualizados e definidos os *trackers* ativos para a presente iteração, é então a altura de os prever ou corrigir. Nesta última fase, o filtro de Kalman de cada *tracker* é atualizado de acordo com a posição do centróide associado na fase anterior. Caso esse *tracker* não tenha valores de atualização, derivado à não existência de uma associação com um *blob*, a sua posição é estimada de acordo com o modelo definido.

O modelo utilizado no filtro de Kalman (KF) é do tipo:

$$x_k = Ax_{k-1} + Bu_k + w_{k-1} \quad (3.2)$$

$$z_k = Hx_k + v_k \quad (3.3)$$

Em que $x_k \in \mathbb{R}^n$ é o estado do sistema linear discreto. $z_k \in \mathbb{R}^m$ é a medida do sistema e $x_k \in \mathbb{R}^1$ a entrada de controlo no instante k . w_{k-1} e v_k representam o ruído de processo e medida respetivamente.

De forma geral, o filtro de Kalman (FK) utiliza medições de grandezas ao longo do tempo, que estão sujeitas a ruídos, para gerar resultados que tendam a se aproximar dos valores reais das grandezas medidas. Isto é alcançado pois este assume que os ruídos a que os modelos de estado estão sujeitos são derivados de distribuições Gaussianas [19]. Desta forma, o FK consegue estimar o valor de uma variável não observável a partir de medições anteriores da mesma. As suas estimativas são calculadas em dois momentos e por isso designadas de *a priori* e *a posteriori*. A primeira estimação x_k^- é feita em $(k-1)$ e a segunda estimação atualiza a primeira aquando do cálculo da variável z_k .

Assim sendo x_k^- e x_k^+ traduzem médias das estimativas, e P_k^- e P_k^+ as suas respetivas covariâncias, determinadas a partir do erro de estimação.

As matrizes P_k^- e P_k^+ traduzem uma indicação de precisão das estimativas do KF. Quanto maiores forem os valores das covariâncias, menor será o grau de confiança na estimação do estado.

A primeira fase no processo de funcionamento do KF é a inicialização, onde devido à sua distribuição Gaussiana, são necessários os parâmetros x_0^+ e P_0^+ , que respetivamente representam a estimativa inicial e a sua incerteza, para dar início à segunda fase: - a previsão. Nesta fase, a estimação do próximo estado do sistema x_k^- (*a priori*) é realizada com base em x_{k-1}^+ . A qualquer estimação existe uma incerteza P_k^- associada, que nesta fase do processo sofre um aumento derivado da evolução do sistema, que está sujeito ao ruído do processo Q_{k-1} . Podem ser observadas em 3.4 e 3.5 as equações que caracterizam esta fase.

$$x_k^- = A_k x_{k-1} + B_k u_k \quad (3.4)$$

$$P_k^- = A_k P_{k-1}^+ A_k^T + Q_{k-1} \quad (3.5)$$

A terceira e última fase do processo é a correção, onde é feita uma atualização da estimativa obtida na fase precedente. Recorrendo à diferença entre a estimação *a priori* do sistema x_k^- e da medição da saída z_k , obtém-se o modelo de *inovação*, o qual se usa na estimativa *a posteriori* (3.9). O ganho de Kalman (3.9), que gera estimativas de mínimo erro quadrático, expressa a quantidade de *inovação* a ser incorporada na próxima fase de correção.

$$\tilde{Y}_k = z_k - H_k x_k^- \quad (3.6)$$

$$S_k = H_k P_k^- H_k^T + R_k \quad (3.7)$$

$$K_k = P_k^- H_k^T (S_k)^{-1} \quad (3.8)$$

$$x_k^+ = x_k^- + K_k * \tilde{Y}_k \quad (3.9)$$

$$P_k^+ = (I - H * K_k) * P_k^- \quad (3.10)$$

Assim sendo, com a aplicação deste método e dos processos anteriores que o alimentam, é possível realizar um seguimento dos objetos pretendidos.

3.4 Resultados

3.4.1 Caso de Estudo

Para comprovar o correto funcionamento do algoritmo, é necessário criar um método de avaliação, bem como analisar o seu desempenho e características comportamentais. Para atingir este fim, foram realizadas diversas experiências dentro do contexto desta dissertação, que se traduzem no seguimento e estimação do movimento de vários objetos. Todos os resultados apresentados nesta secção foram obtidos com um computador com um processador Q9000 2 GHz, e sem qualquer tipo de processamento paralelo ou gráfico.

À semelhança do resultados parciais, apresentados na secção anterior (3.3), a posição estimada pelo sistema de deteção é representada pelas esferas amarelas, e o centróide calculado por cubos azuis. O *output* do sistema de seguimento, em particular a estimação da posição de cada objeto gerado por cada filtro de Kalman, é representado por um cubo verde. O método de avaliação do sistema consistiu na realização de testes em ambiente real que podem dividir-se em dois casos de estudo: um em ambiente interior e outro no exterior.

- **Seguimento de objetos em ambiente interior**

Neste caso de estudo, pretende-se, através da realização de diversas experiências, avaliar o desempenho do sistema em ambientes de luminosidade constante (proveniente de lâmpadas fluorescentes e dispostas em sequência) e de elevada aglomeração de objetos. Com isto, pretendeu-se avaliar o grau de distinção dos objetos detetados no mundo através da combinação de informações obtidas por diversos sensores, isto é, a sua capacidade de, através dos dados fornecidos pelo sistema de deteção, conseguir isolar e extrair a posição do objeto que foi identificado e não dos restantes que se encontram nas proximidades. Para além desta avaliação, pretende-se também analisar as vantagens inerentes à utilização do módulo de seguimento, nomeadamente o modelo do filtro de Kalman, implementado quanto à sua capacidade de previsão de movimentos em cenários que incluam diversos objetos de interesse.

A experiência realizada para alcançar estes objetivos consistiu na movimentação de dois objetos, nomeadamente, a movimentação de duas pessoas, cada uma possuindo um marcador ArUco com identificadores distintos, num espaço fechado, rodeado com uma significativa quantidade de mobiliário de escritório. Os resultados desta experiência são apresentados nas figuras 3.18 a 3.25, onde as pessoas identificadas pelo marcador vão alterando as suas posições ao longo do tempo, realizando movimentos de aproximação e afastamento da câmara não coordenados, nunca ocorrendo cruzamentos nem oclusões.

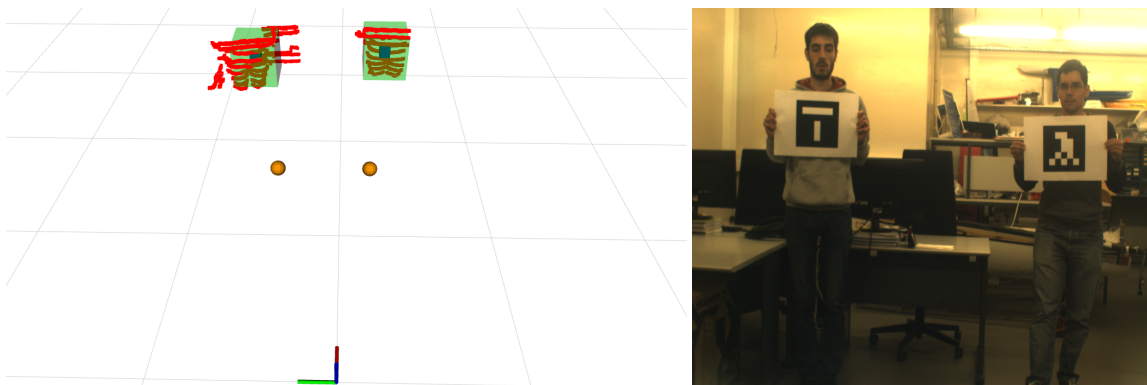


Figura 3.18: Caso de estudo em ambiente fechado.
Frame em $t = 5.4$ s

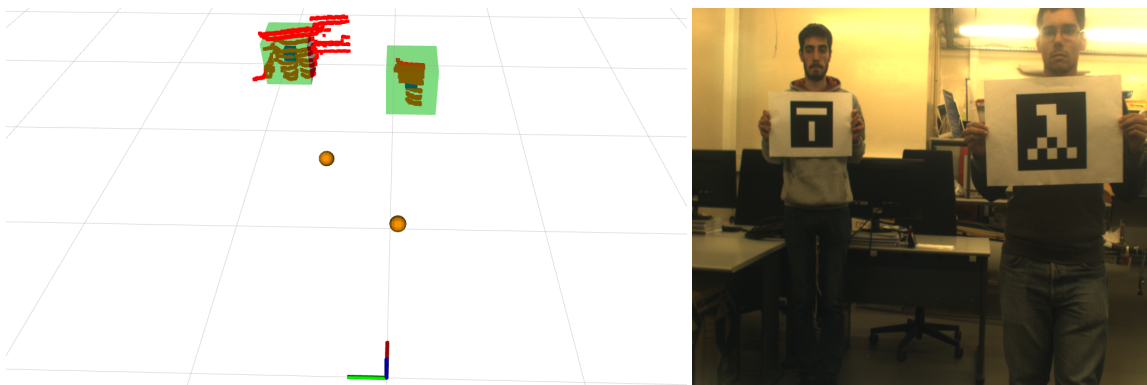


Figura 3.19: Caso de estudo em ambiente fechado.
Frame em $t = 12.5$ s

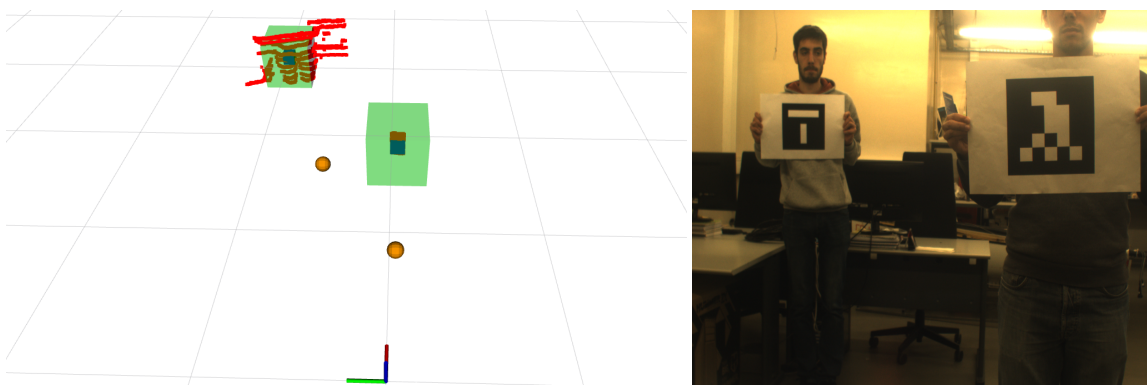


Figura 3.20: Caso de estudo em ambiente fechado.
Frame em $t = 16.1$ s

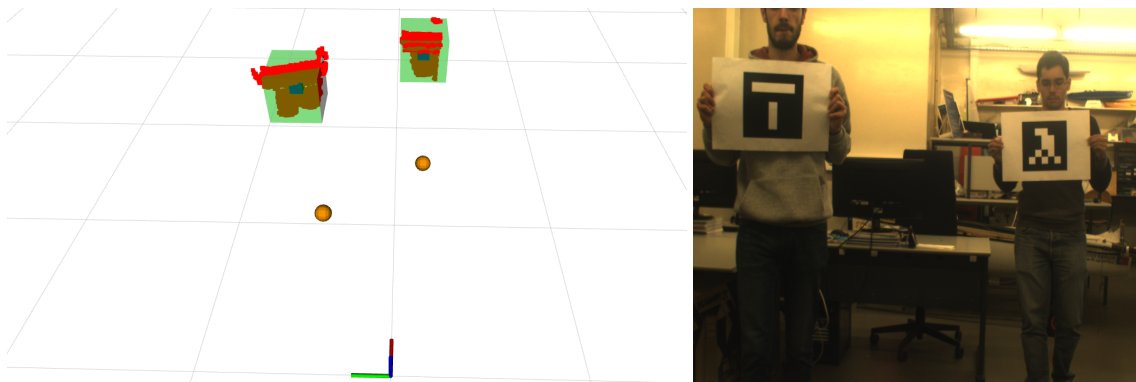


Figura 3.21: Caso de estudo em ambiente fechado.
Frame em $t = 21.7$ s

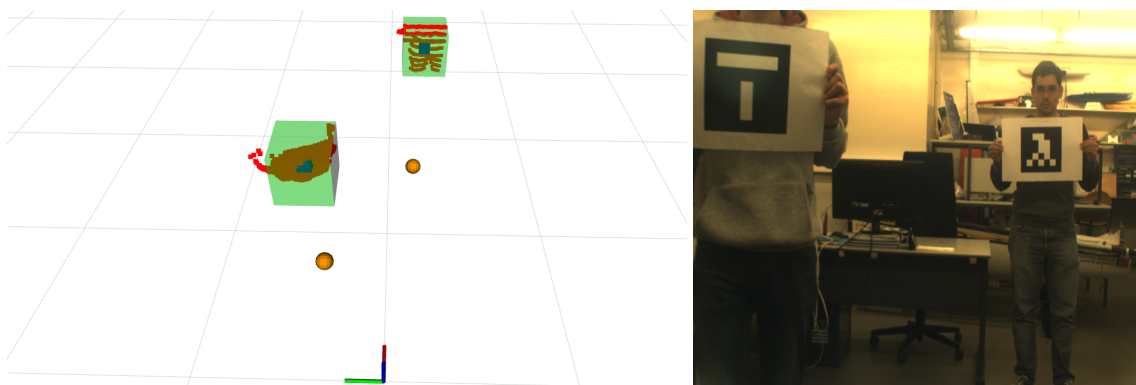


Figura 3.22: Caso de estudo em ambiente fechado.
Frame em $t = 26.3$ s

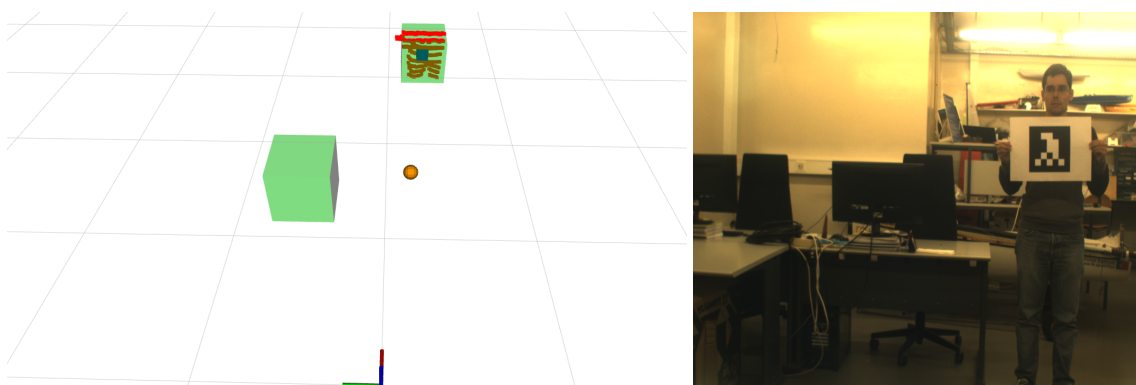


Figura 3.23: Caso de estudo em ambiente fechado.
Frame em $t = 28.2$ s

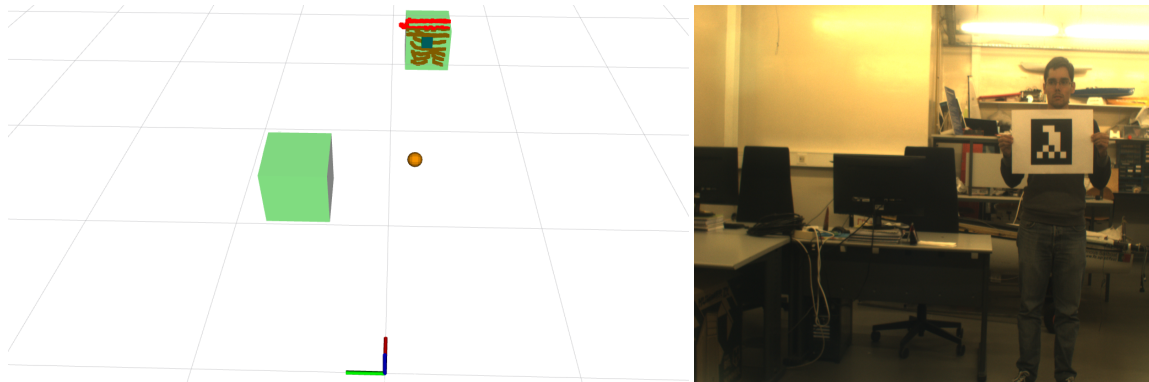


Figura 3.24: Caso de estudo em ambiente fechado.
Frame em $t = 30.7$ s

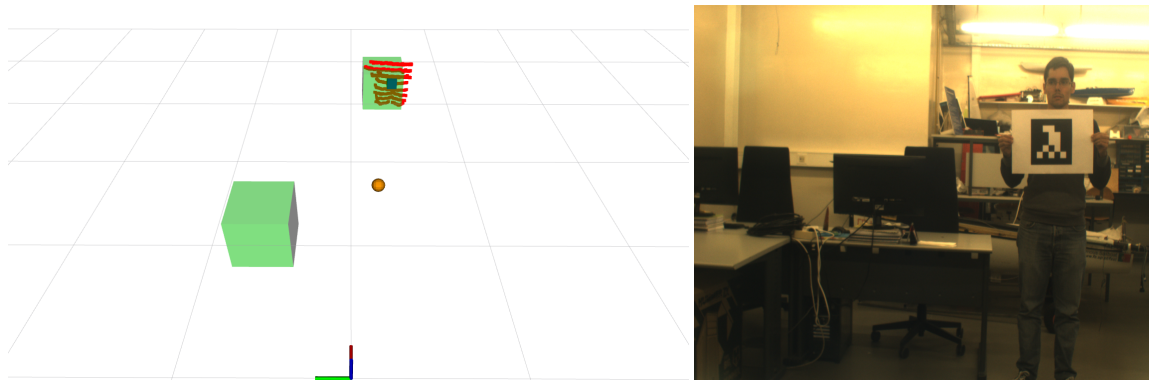


Figura 3.25: Caso de estudo em ambiente fechado.
Frame em $t = 35.5$ s

Nestas sequência de imagens, é possível observar que sistema consegue isolar e extrair, com sucesso, as posições tanto do indivíduo mais à direita, identificado pelo marcador PB, como do indivíduo da esquerda correspondente ao marcador PA que se encontra numa posição de proximidade com outros objetos. Podemos ainda observar o contínuo seguimento, representado pelo cubo verde, dos objetos pelo módulo de seguimento durante as diversas movimentações dos mesmos. Cada cubo corresponde à previsão ou correção da posição que cada filtro de Kalman calcula para o objeto que tem associado. Na figura 3.23, é possível observar o sistema a detetar e a executar o processo de previsão/correção do objeto PB, assim como estimar o movimento do objeto PA que, deixou de ser detetado, uma vez que saiu do campo de visão da câmara, com base no seu perfil de movimento, como pode ser observado ao longo das figuras 3.24 e 3.25. O percurso efetuado pelos dois objetos pode ser observado na figura 3.26, representado pelas linhas contínuas. Na mesma figura, representado pelas linhas a tracejado, é evidenciada a precisão/correção realizada pelo filtro de Kalman respetivo de cada objeto. É de notar que a previsão de posição do marcador PA, realizada pelo sistema, descrita anteriormente, é observada na linha tracejada (verde), após o ponto considerado como fim do percurso. A sua posição é estimada, de acordo com o modelo de Kalman implementado até ao *tracker* ser eliminado.

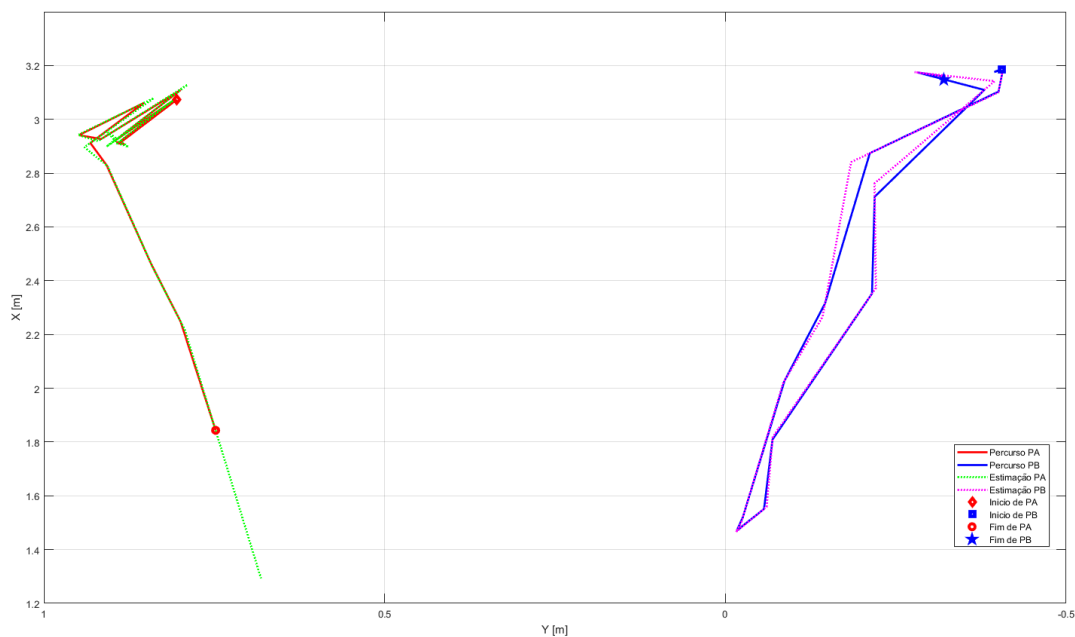


Figura 3.26: Percurso efetuado pelos objetos e o processo de estimação de cada filtro de Kalman em ambiente interior

Através da análise dos resultados desta experiência, pode-se afirmar que o sistema apresenta robustez de extração e determinação da posição dos objetos em ambientes com elevada concentração de corpos, que poderiam induzir erros neste processo. Admite-se ainda, após a análise da figura 3.26, que o modelo do filtro de Kalman implementado apresenta um funcionamento correto e com um elevado valor de certeza quanto às suas estimações.

- **Seguimento de objetos no exterior**

Este segundo caso de estudo teve como objetivo, para além de analisar o funcionamento do sistema num ambiente mais próximo da realidade na qual será usado, em particular o ambiente marítimo, teve também a função de caracterizar o desempenho do sistema, numa situação específica, no seguimento de objetos, a oclusão. Sendo a oclusão uma situação muito comum quando ocorrem mudanças de posição de vários objetos, torna-se importante avaliar a capacidade do sistema de lidar com estes fenómenos e realizar um seguimento conciso e fiável. Neste cenário, o módulo de deteção visual do sistema está sujeito a diferentes condições de luminosidade, sombras, reflexões e efeitos de movimento dos objetos a detetar. Todos estes fatores visuais acabaram por dificultar o correto funcionamento do sistema de deteção escolhido. Apesar de o detetor de marcadores ArUco apresentar resultados positivos, no primeiro caso de estudo, este apresenta uma taxa, de deteção muito baixa em ambiente exterior, na ordem dos 30%. À semelhança do caso de estudo anterior, pode-se, ao longo das figuras 3.27 a 3.32, ordenadas temporalmente, observar os resultados da implementação da arquitetura de seguimento.

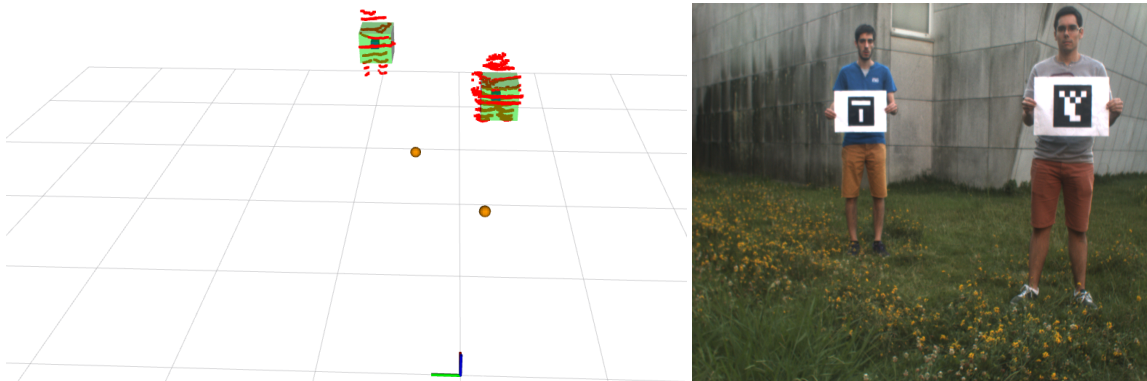


Figura 3.27: Caso de estudo no exterior.

Frame em $t = 5.5$ s

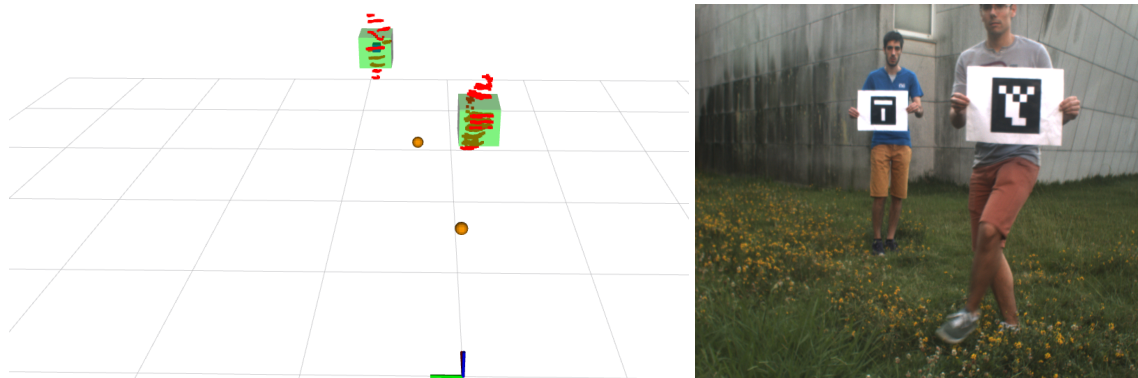


Figura 3.28: Caso de estudo no exterior.
Frame em $t = 9.4$ s

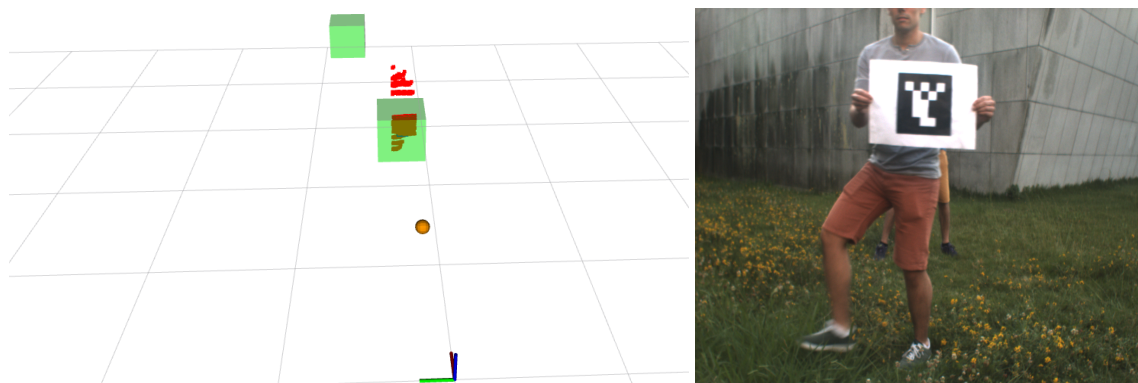


Figura 3.29: Caso de estudo no exterior.
Frame em $t = 10.2$ s

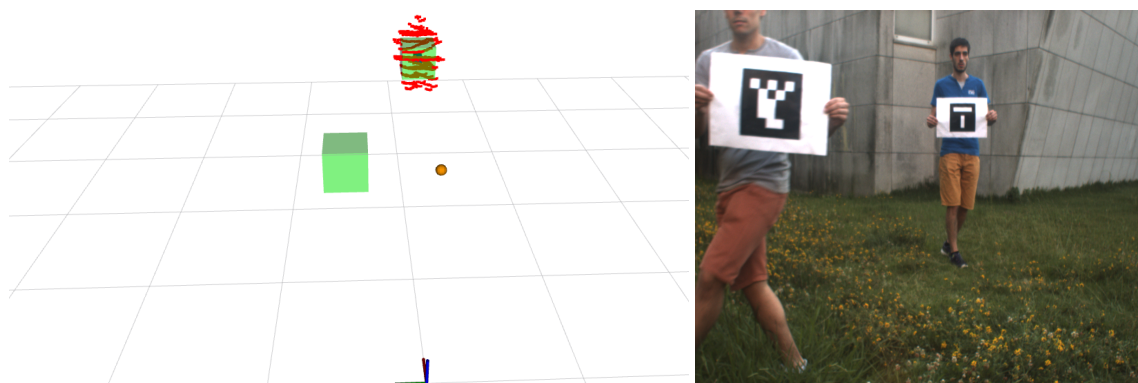


Figura 3.30: Caso de estudo no exterior.
Frame em $t = 13.3$ s

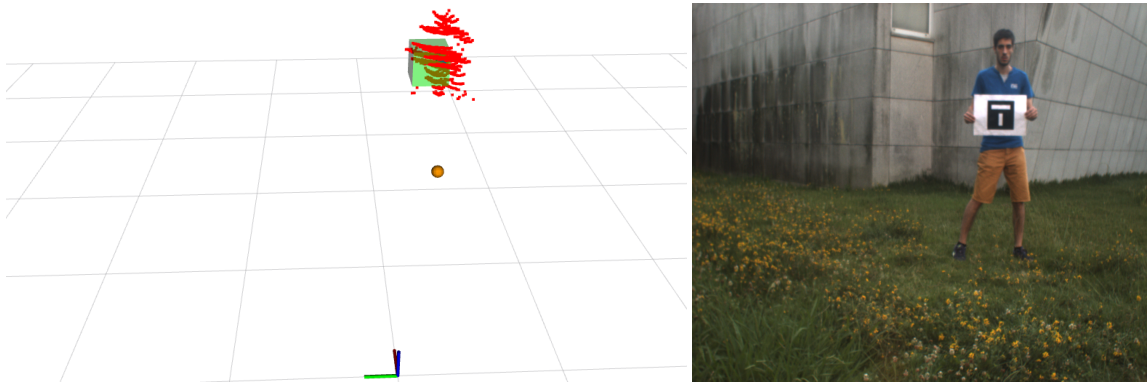


Figura 3.31: Caso de estudo no exterior.
Frame em $t = 15.2$ s

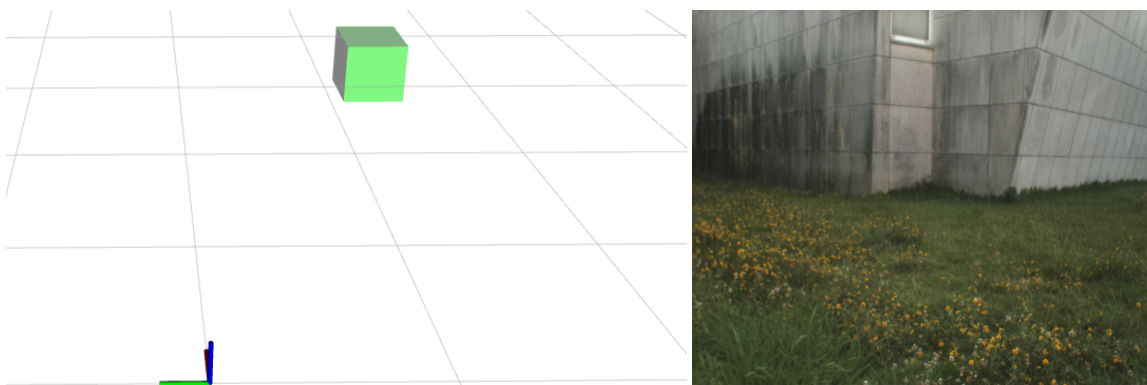


Figura 3.32: Caso de estudo no exterior.
Frame em $t = 19.6$ s

Neste contexto, dois corpos humanos trocam de posição num ambiente exterior, e como pode ser visto na figura 3.29, à passagem do corpo mais próximo da câmara dá-se uma oclusão do corpo mais distante. À semelhança do caso de estudo anterior, encontra-se representado na figura 3.33 o percurso efetuado pelas duas pessoas, assim como a identificação do momento de início e fim da oclusão referida. Através da análise da sequência de figuras, antes, depois e no momento da oclusão, constata-se que o sistema consegue associar a nova deteção do objeto ocluído ao *tracker* criado anteriormente. Primeiro a criação do *tracker*, seguida da oclusão, estimação da posição e por fim, a associação do novo *cluster* detetado ao *tracker* existente, continuando assim, o processo de seguimento. Através da realização desta experiência, pode-se confirmar a robustez do algoritmo relativamente à análise do perfil de movimento dos objetos seguidos e do método de comparação e associação dos mesmos aos *trackers* existentes, reforçando a operabilidade e eficácia deste sistema em diversas condições de seguimento múltiplo.

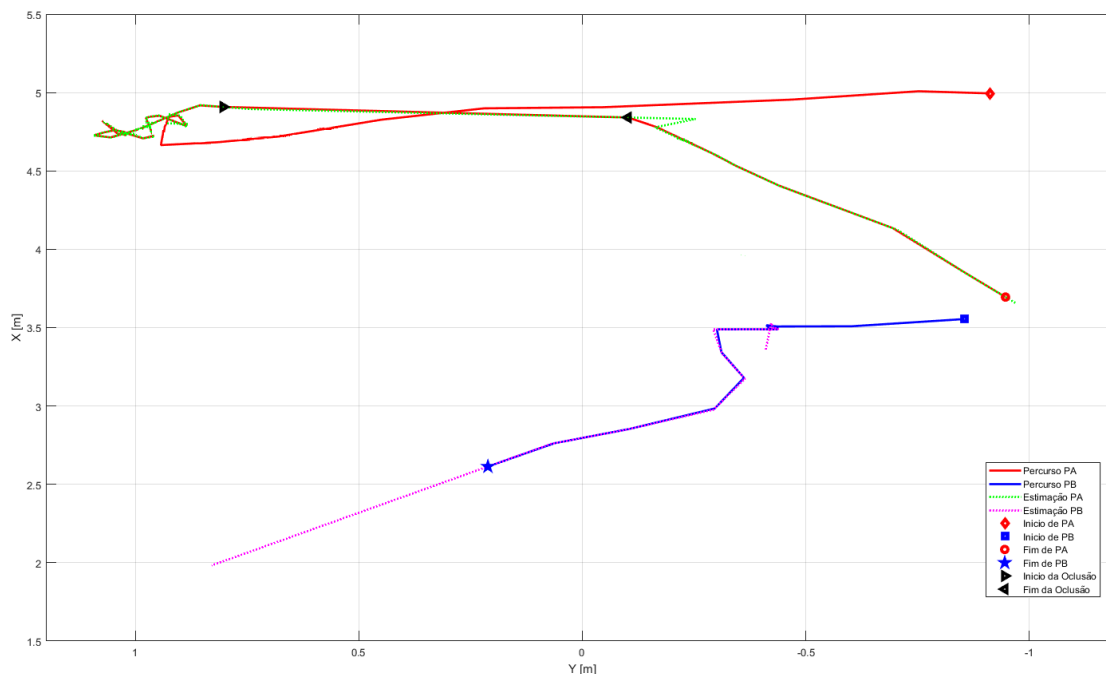


Figura 3.33: Percurso efetuado pelos objetos e o processo de estimação de cada filtro de Kalman em ambiente exterior

3.4.2 Conclusão

Serviu o presente capítulo para descrever a implementação do sistema desenvolvido, não só relativamente às suas especificidades, mas também apresentar os seus resultados parciais que validaram o algoritmo implementado. Com esta validação, foi possível obter resultados a partir de dados recolhidos em ambiente real, mas não marítimo, isto é, com dados adquiridos pelo sistema existente no laboratório do CRAS, no pólo da FEUP.

Inicialmente, foi apresentado o método de deteção escolhido, nomeadamente, o seu processo de funcionamento, as vantagens e desvantagens do seu uso nesta aplicação. É de realçar que o uso dos marcadores codificados do ArUco facilita a deteção, em tempo real, e ajuda, graças aos seus identificadores únicos, o processo de *debugging* do algoritmo. De seguida, procedeu-se ao método de extração mais exata das posições dos objetos detetados, a partir da fusão dos dados dos sensores óticos e do LIDAR. Este processo consistiu na aplicação de múltiplas filtrações e segmentações das nuvens de pontos que o LIDAR consegue fornecer. Nos testes efetuados, esta extração revelou ser eficaz e coerente com as informações transmitidas pelas imagens obtidas. Todavia, o seu tempo de processamento é elevado o suficiente para impedir o funcionamento do sistema em tempo real. No geral, o método de extração tridimensional das posições mostrou-se robusto e eficaz, relativamente às funções que tem que desempenhar. Prontamente, com a certeza do funcionamento dos passos implementados anteriormente, deu-se início ao desenvolvimento e implementação do cerne do algoritmo de seguimento que esta dissertação propõe. De forma a proceder-se a um seguimento e previsão dos movimentos de objetos, foi aplicado um sistema com a capacidade de associar a cada objeto um filtro de Kalman. Estes, através da análise do perfil de movimento dos objetos, conseguem proceder à estimação das suas posições ao longo do tempo.

Por fim, efetuaram-se testes, tanto em ambiente interior como exterior, de forma a testar a robustez e o desempenho do sistema em condições distintas. O algoritmo demonstrou um seguimento correto, coerente e preciso dos objetos detetados. Verificou-se, ainda, o elevado desempenho apresentado pelo modelo dos filtros de Kalman implementados, sendo de realçar a precisão e coesão que este apresentou em situações de oclusão dos objetos seguidos, prevendo as suas trajetórias adequadamente.

Com isto, é possível concluir que o sistema de seguimento, baseado em fusão sensorial, construído, demonstrou aptidão na estimação e seguimento dos movimentos de objetos, tanto a nível de ambiente interior como exterior.

Capítulo 4

Conclusão e Trabalho Futuro

4.1 Conclusão

O uso de veículos autónomos, nomeadamente os ASV, tem sofrido uma procura crescente, quando se necessita de ferramentas capazes de realizar tarefas potencialmente perigosas para o Homem. Esta procura crescente leva, naturalmente, ao desenvolvimento destes sistemas nas mais variadas vertentes, nomeadamente, a perceção constante do ambiente circundante, recorrendo-se, muitas vezes, a métodos de deteção e seguimento. Inerente a este facto, surge a necessidade de dotar estes sistemas com a capacidade de detetar e seguir os objetos à sua volta. Com isto, a utilização de tecnologias passivas e ativas, nomeadamente os sensores óticos e os LIDAR, tem crescido acentuadamente, graças ao seu custo apelativo e ao elevado grau de precisão que estes oferecem atualmente. Desta forma, esta dissertação procurou conjugar e implementar métodos capazes de detetar e, acima de tudo, realizar um seguimento constante sobre uma diversidade de objetos.

Sendo o objetivo a criação de um sistema de seguimento, inicialmente foi necessário selecionar e implementar um sistema de deteção capaz de localizar os diversos objetos no mundo circundante. Sendo o ambiente marítimo o cenário de aplicação deste sistema, a escolha do sistema de deteção recaiu sobre um sistema capaz de detetar simultaneamente e em tempo real, múltiplos objetos, o ArUco. Este sistema de deteção recorre-se de sensores óticos e ao processamento de imagem para estimar as posições tridimensionais dos objetos detetados no mundo. O ArUco tem a vantagem, de através do uso de marcadores codificados, para além de facilitar a identificação dos objetos no mundo, poder ainda realizar esta deteção em tempo real. Apesar deste sistema funcionar como esperado, torna-se necessário, para um preciso processo de seguimento, uma obtenção mais precisa da posição dos objetos. Para isso, recorreu-se à fusão sensorial, nomeadamente, ao uso da tecnologia LIDAR para, com uma maior precisão, adquirir essas posições. Com a determinação mais exata deste objetos, procedeu-se, então, ao seguimento dos mesmos. Este seguimento traduz-se na estimação das suas posições, a partir da criação de perfis de movimento de cada um. Esta estimação e correção das posições foi alcançada através do uso de um filtro de Kalman associado a cada objeto, alvo de seguimento. Todos estes filtros são, então, geridos a um nível superior, permitindo

um correto e robusto funcionamento deste sistema que viabiliza um seguimento simultâneo sobre vários objetos distintos.

Sobre este sistema foram realizados diversos testes com dados reais, isto é, o sistema foi avaliado quanto ao seu desempenho e capacidade de seguimento de objetos em ambientes fechados e ao ar livre. Este demonstrou robustez e um elevado desempenho sempre que se proporcionaram condições favoráveis ao sistema de detecção. Este sistema acabou por ser desfavorável, pois apresentou elevadas dificuldades de funcionamento em ambientes ao ar livre. De forma díspar, o sistema responsável pela gestão dos filtros de Kalman, e que aplica o seguimento, apresentou uma elevada *performance*, relativamente à sua capacidade preditiva, apresentando uma estimação bastante assertiva, mesmo em situações de oclusão.

Em jeito de conclusão, verificou-se que o estudo da literatura e análise comparativa dos métodos de seguimento tiveram uma pesada influência na execução deste projeto, na medida em que influenciaram a tomada de decisões e escolha de métodos a usar a fim de alcançar os resultados finais. É de referir, ainda, que o projeto desenvolvido foi de encontro a quase todas as metas previstas, nomeadamente, ao desenvolvimento de um sistema de seguimento de múltiplos objetos capazes de estimar trajetórias e colocá-las ao dispor do utilizador. A única meta que não foi alcançada foi a criação de um sistema capaz de operar em tempo real. Isto deve-se, como foi referido anteriormente, ao elevado tempo de processamento das nuvens de pontos. Ainda assim, esta dissertação apresenta um método de extração e seguimento tridimensional de objetos que pode ser aplicado nas mais diversas aplicações robóticas.

4.2 Satisfação dos Objetivos e Trabalho Futuro

Tendo em conta o trabalho desenvolvido nesta dissertação, é possível afirmar que todos os objetivos propostos inicialmente foram atingidos. Apesar desta realidade, existe sempre espaço para melhorias, nomeadamente, em termos de eficiência computacional e robustez do algoritmo, de forma a melhorar a sua utilização. De forma a alcançar estas metas, sugerem-se as seguintes tarefas:

- melhorar a eficiência do processo de filtragem da nuvem de pontos, de forma a diminuir o seu tempo de processamento;
- melhorar o *hardware* responsável pelo processamento central do sistema, de modo a diminuir os tempos de processamento e aumentar assim o desempenho;
- reestruturar o algoritmo de segmentação, de forma a aproveitar os 360° de campo de visão do LIDAR e, assim, realizar um seguimento mais abrangente e com menor dependência de imagens; e
- utilizar um sistema de detecção mais robusto e fiável que permita a operacionalidade do sistema num maior número de ambientes e condições.

Referências

- [1] Giancarmine Fasano, Domenico Accardo, Anna Elena Tirri, Antonio Moccia, e Ettore De Lellis. Morphological filtering and target tracking for vision-based UAS sense and avoid. Em *2014 International Conference on Unmanned Aircraft Systems, ICUAS 2014 - Conference Proceedings*, páginas 430–440, 2014. doi:10.1109/ICUAS.2014.6842283.
- [2] Fawzi Nashashibi e Alexandre Bargeton. Laser-based vehicles tracking and classification using occlusion reasoning and confidence estimation. Em *IEEE Intelligent Vehicles Symposium, Proceedings*, número 1, páginas 847–852, 2008. doi:10.1109/IVS.2008.4621244.
- [3] Donald J. Natale, Richard L. Tutwiler, Matthew S. Baran, e John R. Durkin. Using full motion 3D flash LIDAR video for target detection, segmentation, and tracking. Em *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, páginas 21–24, 2010. doi:10.1109/SSIAI.2010.5483929.
- [4] Ahmad K. Aijazi, Paul Checchin, e Laurent Trassoudaine. Multi sensorial data fusion for efficient detection and tracking of road obstacles for inter-distance and anti-collision safety management. Em *2017 3rd International Conference on Control, Automation and Robotics, ICCAR 2017*, páginas 617–621, 2017. doi:10.1109/ICCAR.2017.7942771.
- [5] Allied Vision e Camera Line. High-performance machine vision cameras Better vision for your application.
- [6] Velodyne. Overview. *Http://Velodynelidar.Com/*, página 2800, 2013. URL: <http://velodynelidar.com/>.
- [7] Open Source Computer Vision Library. Detection of ArUco Markers, 2015. URL: https://docs.opencv.org/3.1.0/d5/dae/tutorial_aruco_detection.html.
- [8] Andry Maykol Pinto, Paulo G. Costa, e A. Paulo Moreira. An architecture for visual motion perception of a surveillance-based autonomous robot. *2014 IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2014*, páginas 205–211, 2014. doi:10.1109/ICARSC.2014.6849787.
- [9] S. S. Blackman e R. Popoli. *Design and analysis of modern tracking systems*. Artech House, Norwood MA, 1999.
- [10] Kenji Okuma, Ali Taleghani, Nando de Freitas, James J. Little, e David G. Lowe. A Boosted Particle Filter: Multitarget Detection and Tracking. páginas 28–39. 2004. URL: http://link.springer.com/10.1007/978-3-540-24670-1_3, doi:10.1007/978-3-540-24670-1_3.

- [11] P. Viola e M. Jones. Rapid object detection using a boosted cascade of simple features. Em *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, páginas I–511–I–518. URL: <http://ieeexplore.ieee.org/document/990517/>, doi:10.1109/CVPR.2001.990517.
- [12] Junliang Xing, Haizhou Ai, e Shihong Lao. 2009_CVPR_Multi-Object Tracking through Occlusion by local Tracklets Filtering and Global Tracklets Association with Detection Responses.pdf.
- [13] Horațiu George Todoran e Markus Bader. Extended kalman filter (EKF)-based local SLAM in dynamic environments: A framework. Em *Advances in Intelligent Systems and Computing*, volume 371, páginas 459–469, 2016. doi:10.1007/978-3-319-21290-6_46.
- [14] Jason M. O’Kane e Jason M. O. Kane. *A gentle introduction to ROS*. Independently published, 2013. URL: <http://www.cse.sc.edu/~jokane/agitr/>, doi:8-25-2014.
- [15] Radu Bogdan Rusu e Steve Cousins. 3D is here: Point Cloud Library (PCL). *Proceedings - IEEE International Conference on Robotics and Automation*, páginas 1 – 4, 2011. URL: <http://pointclouds.org/>, doi:10.1109/ICRA.2011.5980567.
- [16] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, e R. Medina-Carnicer. Generation of fiducial marker dictionaries using Mixed Integer Linear Programming. *Pattern Recognition*, 51(October):481–491, 2016. doi:10.1016/j.patcog.2015.09.023.
- [17] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, e M. J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0031320314000235>, doi:10.1016/j.patcog.2014.01.005.
- [18] Tully Foote. Tf: The transform library. *IEEE Conference on Technologies for Practical Robot Applications, TePRA*, 2013. doi:10.1109/TePRA.2013.6556373.
- [19] Mohinder S Grewal e Angus P Andrews. *Kalman Filtering Kalman Filtering: Theory and Practice Using MATLAB*, Second Edition. páginas 0–471.